Irena SAS modeling macros manual

Jan Ilavsky and Peter R. Jemian, *"Irena: tool suite for modeling and analysis of small-angle scattering"*, Journal of Applied Crystallography, vol. 42 (2009). Please e-mail me, if you need copy.

Version 2.37 (for Igor 6)

Jan Ilavsky

Included Methods	**– authors:**	
Desmearing	–	Pete R. Jemian
Pair Distance Dist. Function	–	Jan Ilavsky, Pete Jemian (regularization)
Fractals model	–	Andrew J. Allen
Least Squares Modeling & other methods	–	Jan Ilavsky
Size distribution	–	Pete R. Jemian (Maximum entropy/regularization)
Unified model	–	Greg Beaucage
Reflectivity (aka Parrat's code)	–	Andrew Nelson

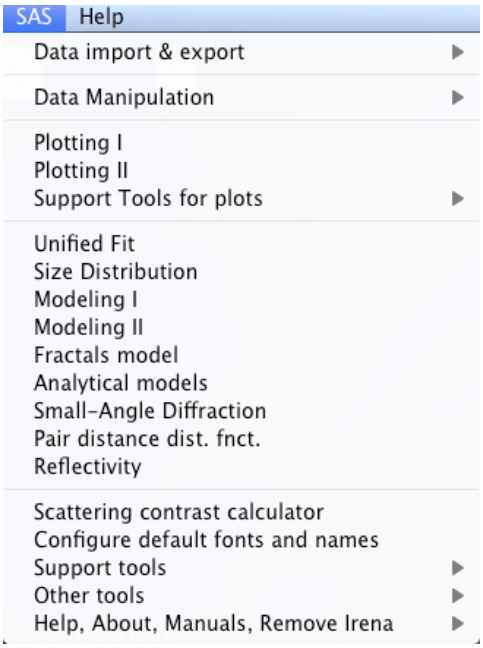
Wednesday, February 24, 2010

Disclaimer:
These macros represent a collaborative work in progress and it is very likely that not all features are finished at any given time. Therefore, some features may not work fully or at all. Please note, while I try my best to verify the results, no guarantees can be made as to the reliability of these results. Please, verify results in some other way. Please report any bugs to me, I will do my best to remove them ASAP. I provide limited support for users of these macros. Limited means that my time available for this support is limited. If you need help, e-mail Igor file to me with data so I can work on your data.
ilavsky@aps.anl.gov

Content

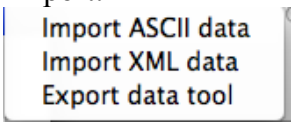
Location of items in SAS menu.

The location of items in SAS menu keeps changing. Here is current (version 2.30) location to help you find what you need…:

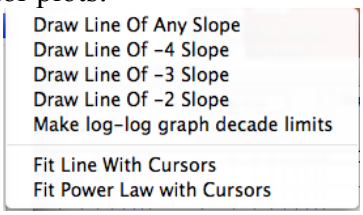


Submenus:

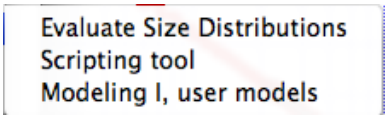
Data Import & Export:



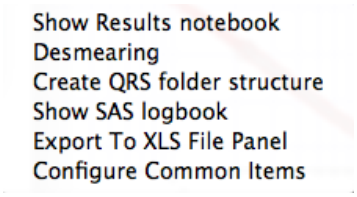
Support tools for plots:



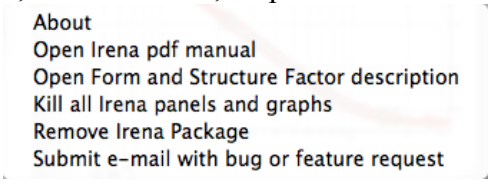
Support tools:



Other tools:



About, Manuals, Remove Irena, help



Introduction

The "Irena" package is a suite of Igor Pro (Wavemetrics, version 6) macros for the evaluation of small-angle scattering data. It has been designed to use seamlessly data from APS USAXS instrument (beamline 32ID, Advanced Photon Source, Argonne, IL; evaluated using "Indra" package). Further it has been designed to work very nicely with data using "qrs" naming convention, but it can be used for any SAS data. Its use for other data is made easy by a customized import tool, which should be able to handle most "column-type" ASCII data from various SAS instruments.
The package contains following parts:

1. **Size distribution** – using Maximum Entropy, Total Non-negative least square (TNNLS) & Regularization methods for evaluation of small-angle scattering from scatterers represented by number of different form factors.
2. **Modeling** (I) of scattering from up to 5 populations of various-shape scatterers, with least-square fitting optimization of model parameters
3. **Modeling** (II) of SAS from up to 6 populations of data to multiple data sets… Complicated, VERY powerful tool. Number of form factors and structure factors.
4. **Unified model** for fitting SAS data using up to 5 levels of combinations of Guinier and power law dependencies
5. **Pair distance distribution function** (PDDF, p(r))
6. **Fractal model** – combination of 2 mass and 2 surface fractals.
7. **Analytical models –** tool with option for:
    - **Debye-Bueche model** for scattering from gels
    - **Treubner_Strey model** for small-angle diffraction
8. **X-ray & neutron reflectivity** calculations using Parrat's recursive method
9. **Scattering contrast calculator** including anomalous effects
10. **Data import tools**. Allows importing multiple ASCII or XML files, where SAS data are written in columns, separated by white space, tab or other separators. Allows creating user friendly logical folder structure within Igor experiment.
11. **Desmearing** for finite-slit length smeared data
12. **Data manipulation tool**. Allows merging, smoothing, adding together and subtracting of SAS data sets. These data sets do not have to necessarily use the same naming convention.
13. **Two plotting tools**. This tool allows to generate various SAS plots (Porod, Guinier, Kratky, Zimm…) and do some basic fitting. Further the tool allows to save plot styles with various formatting parameters and then fast reapply these on other data sets, generating exactly same plots useable for publications.
14. **Data "mining" tool** – allows searching for results (variables/strings/waves …) in the folders of Igor experiment with flexible output options.
15. **Data export tool** – exports into ASCII various types of data.
16. **Scripting tool**. Tool to run (Size distribution and Unified fit for now) tools on multiple data sets at once.
17. Tool to create folder structure for unstructured QRS data. And few other tools…

Methods use as similar as reasonably possible. This should simplify learning curve for the users…

Name "*Irena"* of this package, for those really interested, is the name of my wife. All my packages have female names, for example "*Nika*" is the name of my daughter (Veronika), etc. As choice of names is more or less arbitrary, I felt that selecting the name of my wife for this large and important package will be one way to give her credit for all the time I spent working on this package and not with my family.

Instructions on installation

To install the macros, please install first Igor Pro, version 5 or 6 and update to latest release.
There re two ways to install the macros:

I. Old way (user with administrative privileges)

Copy the macro files from the distribution zip file using the folder structure inside the zip file as guidance on where the files belong. Note, that most files belong into the subfolder "Irena" inside the folder "…Wavemetrics:Igor Pro:User procedures" and at least one file belongs into the folder "…Wavemetrics:Igor Pro:Igor procedures". The files located in "Igor Pro" folder are useful only on PCs and only in some versions of Windows…

II. New way (user with limited privileges)

Place the files in the zip file in any appropriate place where you, the user, has write privileges. Note, that the files are organized in folders. Then create shortcuts (links) in the "…Wavemetrics:Igor Pro Folder:User procedures" to the "Irena" and other folders in the "User procedures" from the zip file (where ever you placed them) and create a shortcut (link) in the "…Wavemetrics:Igor Pro Folder:Igor Procedures" to the "Igor Procedures" from the zip file. YOU WILL NEED ADMINISTRATIVE privileges for creating these shortcuts/links…

Whichever method you choose, the macros should work the same. Note, that when files are created by the macros (compounds, graphs, temp files…) they are stored with the macros themselves and therefore, if user is NOT running as administrator and has not writing privileges where macros are stored, nasty OS errors will apprear!!!

To load macros, **select "Load Irena SAS macros" from "Macros" menu** after starting Igor Pro.

Please, learn more about full capabilities of the Igor Pro. It is very powerful graphing and data evaluation package. It may be necessary for you to handle data import and handling, data export and some graphing. Further, the macros heavily rely on the data folder structure, so it is important to learn enough to realize the use of this feature…

Please read these comments:

Few suggestions first:

1. Learn enough Igor, that Igor problems do not prevent you from getting results. Igor tour and 1-2 hours playing with it should be sufficient
2. Read this manual full and test what is shown on your own computer
3. Use folder structure, or things will become way too messy for these tools to be useful
4. Read supporting literature (especially papers about Unified fit) if you want to use these methods

Comment on ending the macros:

At any time user can end working with the macros by closing associated graphs and panels. There is also command which closes all open windows and panels of this package.

Getting help

There is no real help file written yet for this code. However, most of the panel items do have help associated with them and if you move cursor the help text should be visible in the bar at the bottom of the screen (on

Windows) or in the balloon help (on Mac, if you switch the balloon help on). Please, use it; I tried to be descriptive there.

Important Information

The code uses for all size related parameters Angstroems (10^{-10} m) or for Q vector (A^{-1}). In the case of scattering contrast, number distribution and any other volume contents centimeters (10^{-2} m).
Output of the size is *usually* in particle diameters, but read the graphs, the output may not be always the same. Output graph legend or panel text should be always correct.

Kill all Irena panels and graphs

This menu item allows to close all Irena related windows – panels and graphs – to be closed at once. Very convenient…

Open Irena pdf manual

In most cases this should open Irena manual in default pdf reader. If you are reading this, you probably managed.

Open Form factor description

This should open pdf file with form factors description – including simplified code and graphs. These are form factors in the "central bank" of the Irena, available for use in packages, which use them.

0. GUI controls and common controls

The control panels now share some common controls. Here is summary of common features for these controls:

0.1 Configure default fonts and names

"Configure default fonts and names" in the SAS menu will create panel with some controls common for all tools, like font type & size and how legend names are handled. NOTE: Panel controls are applied immediately to all existing panels, graph controls are applied ONLY to the newly created graphs (aqnd only those wehich were upgraded to this behavior).

Panels font and font sizes

These controls enable user to customize font used on control panels therefore this enables customization for a given platform. This is necessary as more and more control is provided on each platform to user and therefore default fonts and font sizes may not be appropriate any more for the panels I design. These settings are actually saved on a given machine as well as the experiment. This has some interesting features, so please, read carefully:

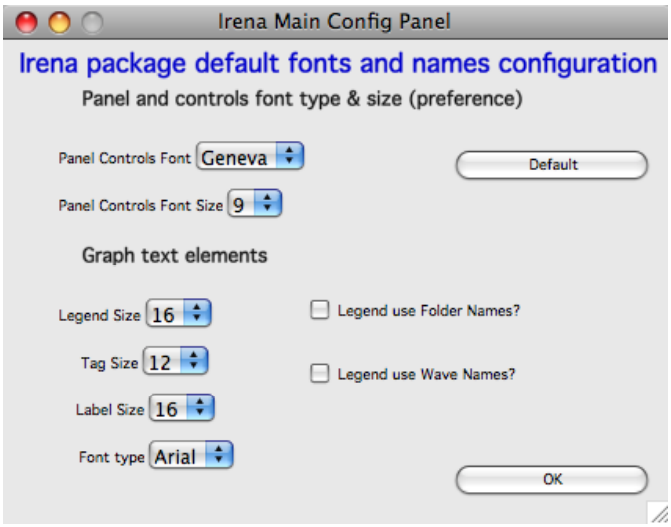
When these controls are run (and user is forced to run them if the Irena is loaded and preferences are not found), they save preferences in special folder Igor maintains for users. At the same time, the settings are applied to the current experiment.

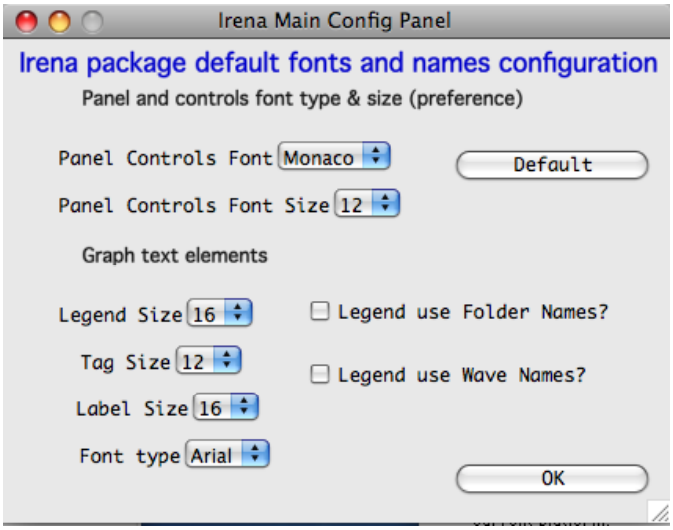
When this experiment is opened on another computer, the preferences from that computer are not reloaded, so the experiment will use preferences from the original computer. When the "Configure GUI and Graph defaults" is run, it will reload the computer defaults and apply them to the given experiment. Then user can change the fonts and font sizes as they wish. The new settings are saved on the computer – and within the experiment.

Note, that Panel font and font size are platform specific, so same experiment may present differently looking panels on Mac and PC.

Note, not all controls actually follow these settings, I have been changing some buttons to specific font and font size and those are not affected by these settings.

If there are any issues with the behavior, please, let me know and I'll see if I can make it more logical.





Note the difference in Configure GUI and Graph defaults panels when different fonts are used. Left is using Geneva font size 9, right is using Monaco size 12, both on Mac platform. You can mess up the panels really well by wrong choices!

Defaults button returns the panel font choices to platform specific default state (Mac: Geneva size 9 and PC Tahoma size 12). Note, that there is no guarantee that these were your choices before. But these should be reasonable choices for most setups.

Graph controls
I am slowly adding in various parts of the whole package calls to these commonly stored values. This allows user to configure fonts for various screen sizes. This seems necessary to allow use of Mac/Win platforms with vastly different screen sizes and resolutions.
Not all packages follow these controls yet, if you see issues in package of your choice, let me know and I will try to address them ASAP. Time is limited resource.

0.2 Data input

Data input part of the panels is served by common package (mostly) and has more or less similar behavior – with modifications appropriate for each package. The purpose of these controls is to provide as much help to user to select appropriate data as possible. This is not easy task… Sometimes even it is not clear what the right help is.
There are few checkboxes for data types, up to 4 popups with Data Folder, Wave with X, Y and error data. If Model input is appropriate, Qmin, Qmax, number of points and log/lin binning inputs are displayed.

How the control works:

Type of data:
Indra 2 data data from Indra package (DSM_Int, etc.). Assumes data are in root:USAXS folder (or any subfolder) only.
QRS data data with q_name, r_name (intensity) and optionally s_name (error).
Model No data, tool will create q data using user input and intensity/error data will be set to 0. Then passed intot he tool so one can model with no measured data present. Available ONLY when appropriate
Irena results should know results from Irena package (all different types). When appropriate will be available. Note, that in any folder may be number of different results available.
User type currently not used, but allows definition of any other naming structure to be used in the future. Note this can be named differently at any time and can provide access to any doublet or triplet of wave types, if it can be defined.
No type of data selected In this case the tool will present choice of all folder in the experiment and for data waves all of the wave in the particular folder. This method will work always, but may be quite challenging to use.
Basic control logic
When particular type of data is selected, the tool should go and find all of the folders containing at least one of the type of data.
Indra 2 data at least one of M_DSM_Int (M_DSM_Qvec, M_DSM_Error), DSM_Int, M_SMR_Int, SMR_Int triplets.
QRS data triplet of waves starting with q, r, s with the rest of name the same. Note, this is the most cpu challenging data type, so it will take the longest.
Irena results any of the results from Irena package. If any is missing, let me know, please…
Model no input data, input data will be created.
User not used at this time. Can be used in the future for any data types which can be defined.
Nothing all folder

These folders are presented in the "Data folder" for user selection. When user selects the folder, rest of Wave popups will be populated by first valid set, which is in the order prescribed by internal logic.

If other data set is needed, select different data in the "Wave with X axis dala" popup. This will attempt to fill the next ones with appropriate data. This may not be unique, so the first match will be filled in.
Then if still necessary, fill in the other two popups.
Note, that it is possible, that depending on tool you can select only two data waves (X and Y), some tools may require also error wave.

0.3 Use of XOP

Igor Pro enables use of external C-code to speed up some high cpu intensive operations. Currently various optional xop program are available:

1. Two by Andrew Nelson http://motofit.sourceforge.net/wiki/index.php/Main_Page – one for calculation of reflectivity (abeles.xop) and one for genetic optimization (GenOpt.xop). Both are optional and need to be placed in "Igor extensions" folder. Both speed up the calculations by factor of two (about) compared to Igor code.
2. XML loader (also by Andrew Nelson) necessary to load XML (CanSAS) file formats. You can download this general use XML xop from : http://www.igorexchange.com/project/XMLutils
3. In the future, Form factor will be provided by xop library maintained by NIST reactor. Once that is available, the link to download place will be provided here.

0.4 Genetic optimization

Genetic optimization method is form of fitting from SAS data. It has been developed for optimization of reflectivity data but is very useful for cases where least square fitting may not find global minimum. It has been programmed for Igor by Andrew Nelson, who is also author of internal code for reflectivity tool.
Note that this code uses some version of Monte Carlo method. Therefore limits are _very_ important. When Genetic optimization method is used (note: IT TAKES LONG to run) user will be presented with dialog to check the limits. For this method is really important that the calculations do not fail for any combination of parameters and that the range of probed parameters is sensible.

1. Loading data – ASCII & XML (CanSAS)

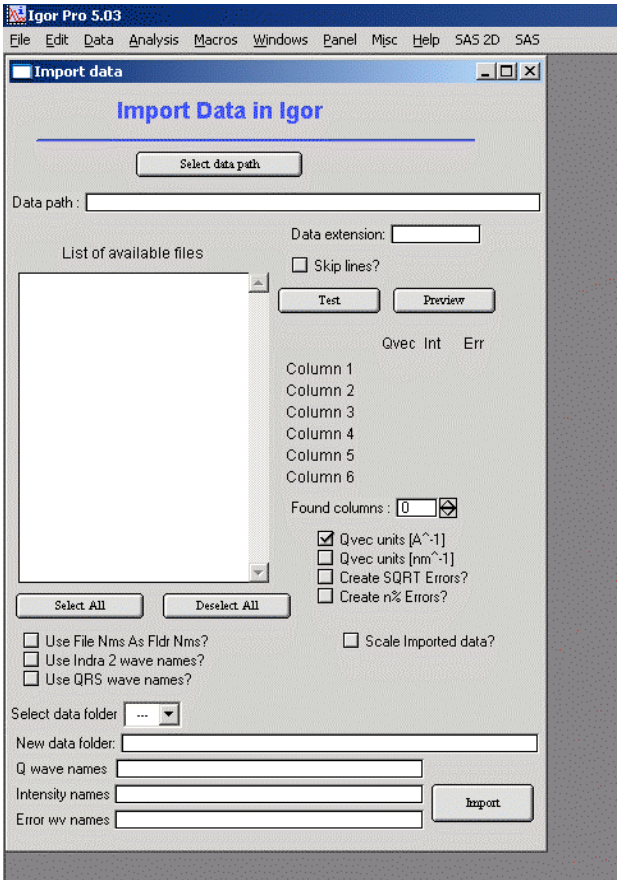
Users of Indra 2 produced data can skip this chapter.
I have included ASCII set of data from our USAXS experiment as example on which user can play and test capabilities.

Comments:
When loading data for use with Irena macros, the user needs to decide ahead on naming system, which will be used. Indra 2 (USAXS) data use somehow specific naming structure: folder names for sample names and within the folders the waves with names such as: DSM_Int for intensity wave, DSM_Qvec for Q vector wave and DSM_Error for error wave. This may not be the first choice of most users from other instruments. Therefore the Irena handles well also "qrs" naming convention, where the wave with q vector starts with q_*andTheDataName*, intensity wave is named r_*andTheDataName*, and error wave s_*andTheDataName*. This allows placing multiple data sets in one Igor folder, **which I strongly discourage**. Note, that when Irena macros save results within Igor experiment, they **ASSUME**, that they can write simply solution into folder where the SAS data came from. Irena macros **WILL NOT** overwrite old results, but it may be impossible to figure out, which data the particular result belongs to…. To make best use of these macros, please use folder structure with folder names being the sample name.

YOU WERE WARNED!!!!

1.1 Description

In "Data import & export" from "SAS" menu select either "Import ASCII data" or "Import XML data". Following screens appears:
1. ASCII



Explanation of control available here:
"*Select data path"* browse to the folder on the computer drive where the data to be imported reside.
"*Data path"* this shows the path selected above. Cannot be edited in this window, use button *Select data path*.
"*List of available files"* lists all files in the current folder on the computer, unless masked by *Data extension*. One or more files here can be selected for import. Use shift - click to select multiple files (on Windows) or whatever – click on Macs.
"*Data extension"* if extension is put in this filed (e.g., "dat") only files with the "dat" extension will be shown in the *List of available files*.
"Skip lines" if there are known number of lines, which need to be skipped. Note, Igor will automatically read file structure and skip usually any text header which needs ot be skipped. Therefore this "skip lines" should not be usually necessary.
"Test" Test import of first selected file. Not really necessary, but very useful. Sets checkboxes for Column 1 to 6, how many columns were found in the file, etc.
"Preview" Opens the first selected file in Igor notebook for preview. Kill notebook after use, it is not needed for anything else…

"column 1 – 6" and *Qvec Int err*" This is checkbox area, in which user needs to select which column of data contains which SAS data. Assumption is, that SAS data are in the first 6 columns in the ASCII file. These checkboxes appear when "*Found columns"* number gets set. User can set it or it gets set during "*test*". "*Select all"* or "*Deselect all"* modifies which files are selected in "*List of available files"*.

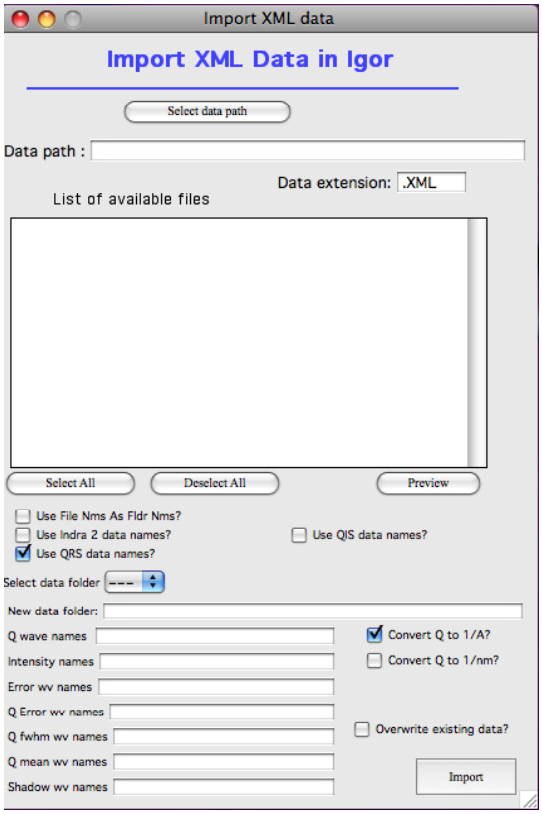
"Qvec units" select proper checkbox. Units will be converted to A^{-1} if nm^{-1} data are imported. Irena uses A^{-1}.
"*Create errors"* if the data imported do not contain error bars, this will generate sqrt(Intensity) error bars. These can be further modified (multiplied) in Data manipulation tool.
"Scale imported data?" if the data need to be scaled by some calibration factor… New input variable appears, if necessary.

"*Use file name as folder name"* **Strongly suggested to use**. Will cause the import tool to create for each imported data set new folder with name by the file name.
"*Use Indra 2 wave names"* and "*Use qrs wave names"* selects which naming structure is used during import of data. One of these selections is more or less necessary for multiple file import. Single file import can be done by manually filling the following controls.

"*Select data folder"* and "*New data folder"* Using pull-down menu in *Select data folder* user can select existing data folder where to put the imported data. Using *New data folder* user can create folder in Igor for the data. Note, that "<filename>" will be replaced with the file name of the imported data file during import. This allows for creating data structure which uses folders during multiple file import.

"*Intensity wv name"*, "*Q wave name"*, and "*Error wave name"* – these can be filled with the names for data waves. Note, that "<filename>" will be replaced with the file name of the imported data file during import.

"*Import"* imports the selected data.

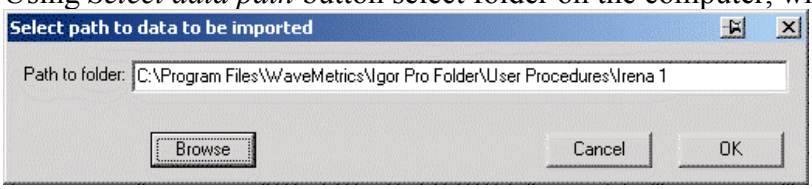


Some of the controls (checkboxes) do change some of the setting in other controls. Generally the proper order, how to select and modify control is from top to bottom.

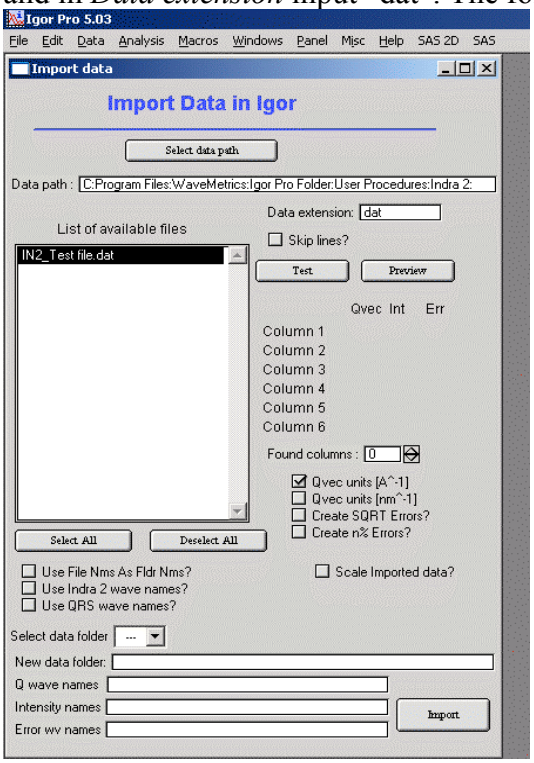
2. XML data file:
NOTE: XML data tool requires xop for XML data file interface. See chapter 0.4 above for the link to this file.
Similar controls, except CanSAS XML file does not need some of the controls. Therefore, the GUI can be easier. On the other hand there may be more data columns (meaningful) in this data file and while Irena does not use any of these, they can be loaded to be useful for user code or other tools, which may be able to use them (like NIST macros).

1.2 Importing test file:

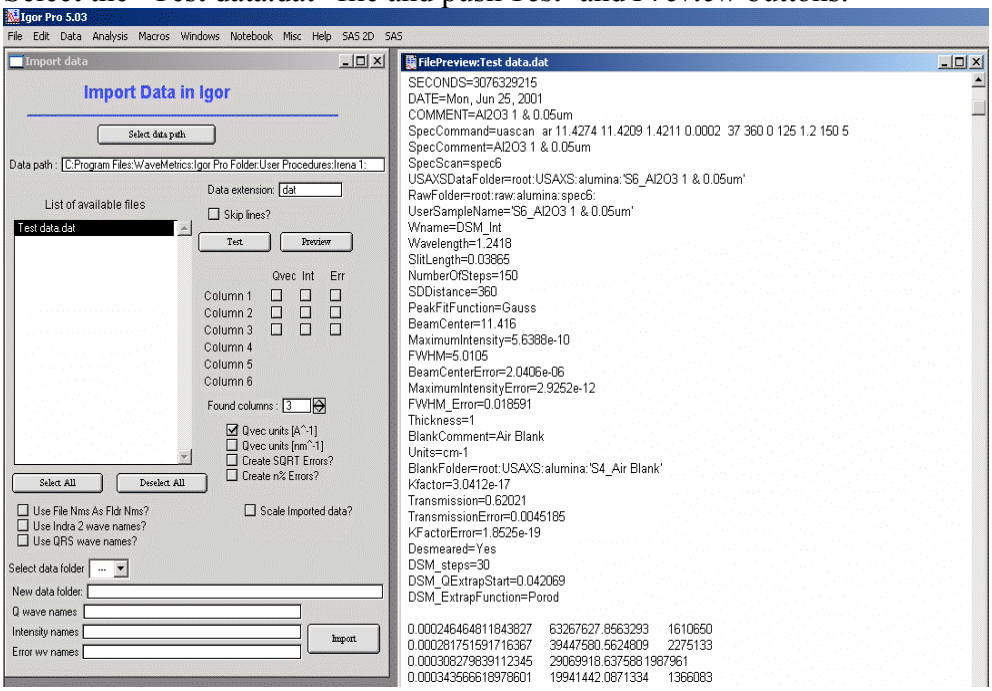
Using *Select data path* button select folder on the computer, where Irena data are installed, for example:



and in *Data extension* input "dat". The following should be the panel:

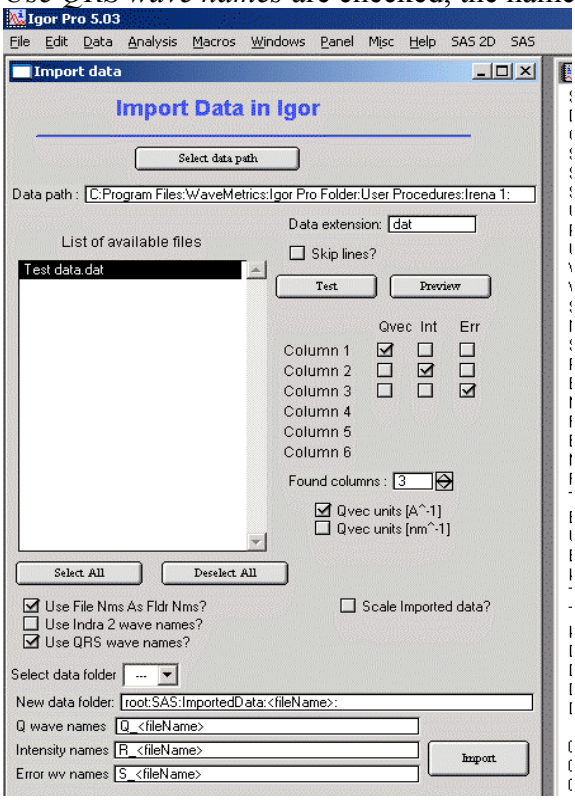


Select the "Test data.dat" file and push *Test* and *Preview* buttons.

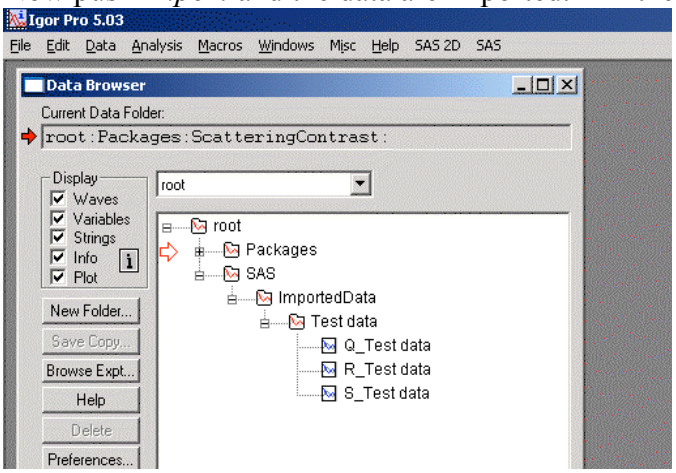


Igor found 3 columns of data so 3 rows of checkboxes appeared. The *Preview* has created notebook on right, where user can preview the file and check, which columns contain which data. Note, that Igor skipped the block of text in the beginning of the data file automatically.

Check cheboxes according to following screen and noticed, that *Create errors* checkbox becomes unavailable when any checkbox in the Err column is selected. Notice, that when checboxes *Use file nms as Fldr Nms* and *Use QRS wave names* are checked, the names for folder and data wave names are filled in with default.



Now push *Import* and the data are imported. Kill the Import data panel and see in Data browser:

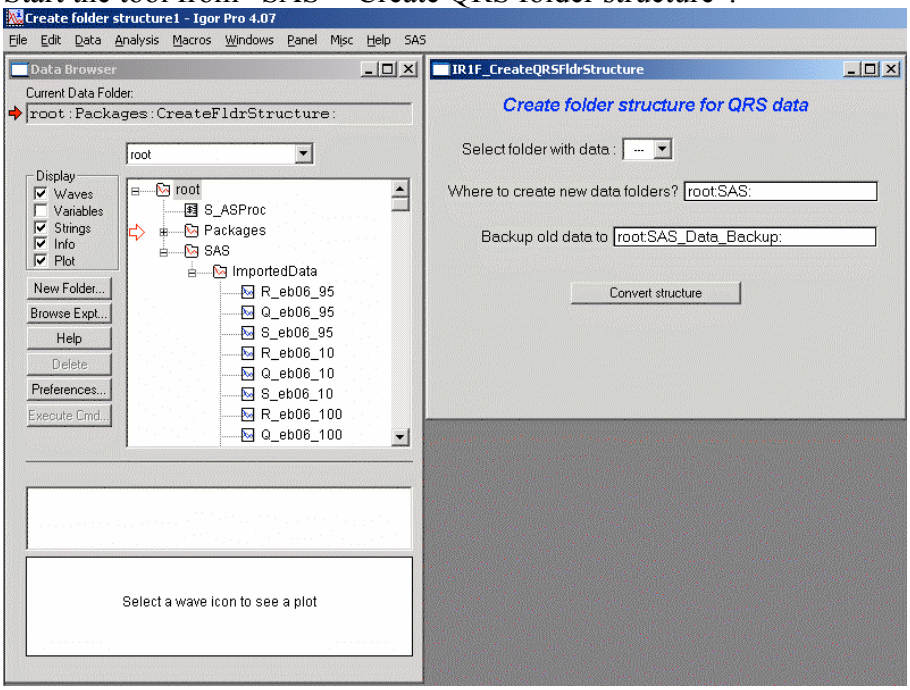


2. QRS data folder creation tool

2.1 Introduction

Many users may have QRS named data in unstructured way – that is all data placed in one folder, very often "root" folder. This is not very convenient place for the data, since the Irena macros make heavy use of the folder structure. To help the users, I made little simple tool, which should in most cases create successfully folder structure for this type of data.

Start the tool from "SAS" "Create QRS folder structure".

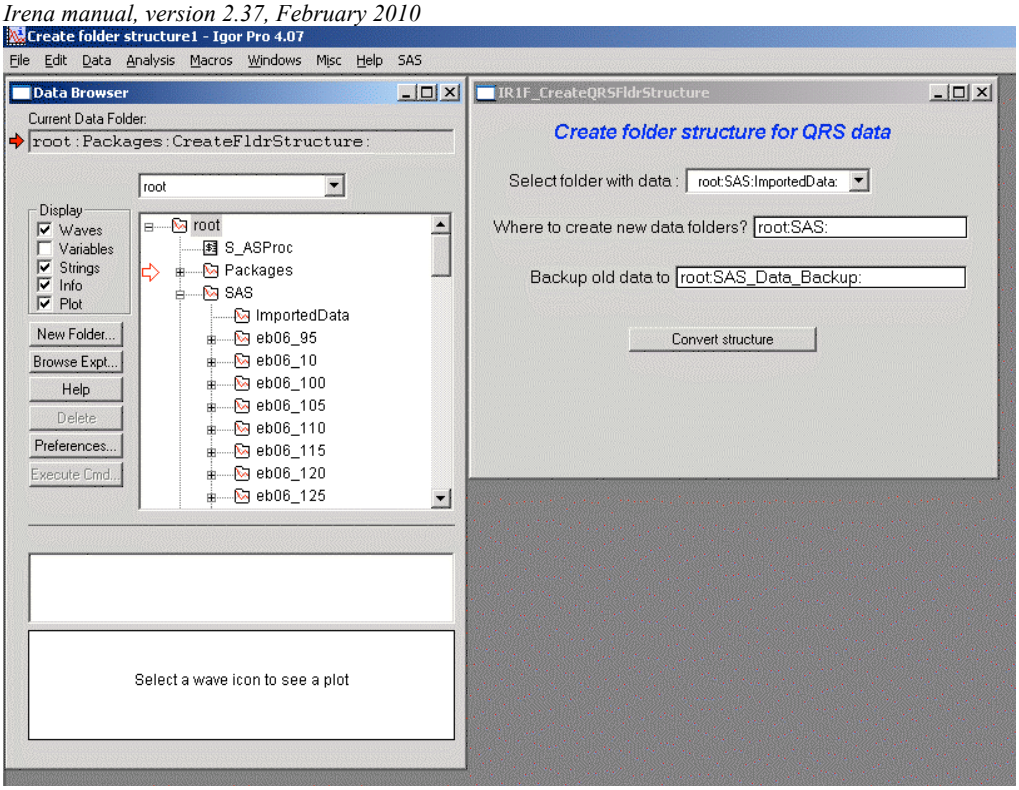


Note that all my imported data are in "root:SAS:ImportedData:" folder. They can be in any folder in the Igor experiment. Note few controls in the panel just created.

"*Select folder with data"* this popup will list ONLY folders containing triplets of QRS named data. Select folder, which contains data you want to convert. In case of this example the "root:SAS:ImportedData:" folder

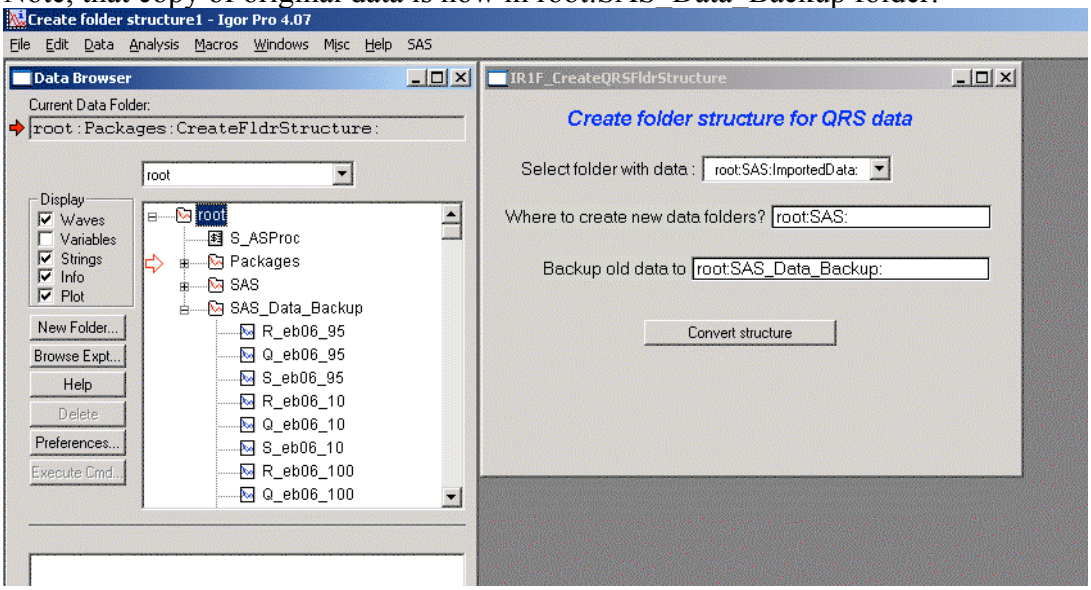
"Where to create new data folder?" Input full folder name to folder, in which you want to create new folders with the separated data

"Backup old data to" input full folder name where you want to put backup copy of old data. If empty, backup will not be created.



Select appropriate folder with data and push button "Convert structure". Result can be seen above – folder "Imported data" is now empty and new folders which are named by sample names (using the name from QRS naming structure) were created. Each contains QRS named triplet of waves.

Note, that copy of original data is now in root:SAS_Data_Backup folder:



Few comments.

This tool is relatively simple and does not do much checking. It will not be able to remove waves, which are part of any graph (or for other reasons Igor refuses to remove them). It will create new copies of these data, it just cannot remove the waves in use.

The folders with data are never overwritten, if folder of the particular name exists, index starting from 0 will be attached to the name.

Do not backup into the same place where the data are coming from. Make separate backup into separate folder.

Data, for which the code does not find properly named QRS triplet of waves are not touched.

There is no checking for wave length or other validity, all what is used is the names of the waves.

The code does not know about any “name extensions”, so data named “R_myName_BkgSub” are treated as separate data from original data “R_myName”…

I assume, that your names are legal and valid. The code may fail on liberal names (names with spaces and other weird characters). I need to test that later. This should not be a problem, since most users with the data needing this treatment should have standard (non-liberal) names, or the code used to create these should not work..

3. Data manipulation tools

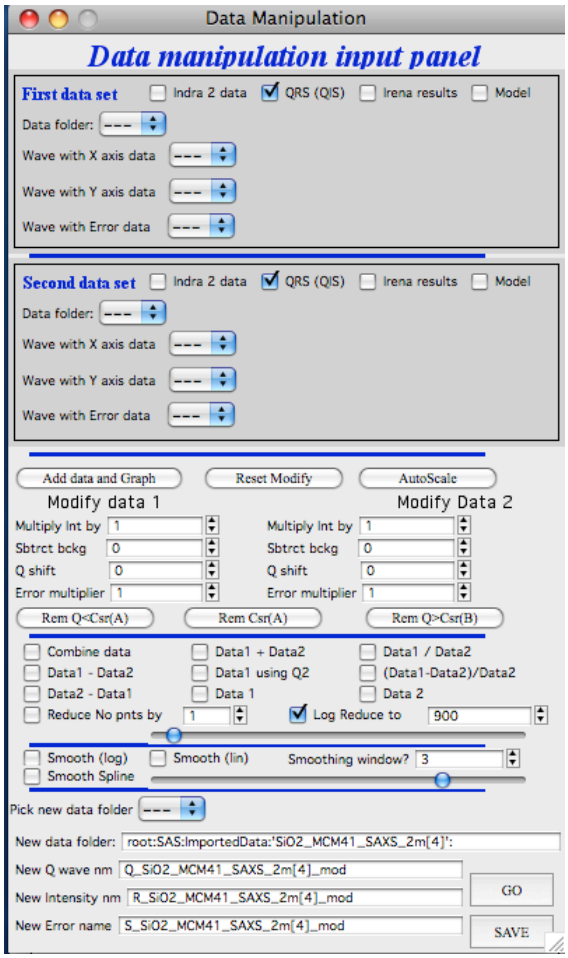
3.1 Data manipulation I – one or two data sets

This tool allows the user to modify data in many ways. It can work on one or two data sets at the same time. The data sets may or may not contain errors.

The tool allows the following data manipulations:

1. Modify data – multiply the intensity and errors by constant, subtract background, and shift in Q
2. Merge data sets - attach data sets together as when attaching sectors from different camera lengths on pinhole cameras. The data can be scaled together manually or automatically by using area under the curves in range selected by cursors
3. Sum together data
4. Subtract data set one from another
5. Re-bin intensity and errors to new Q
6. Smooth data in linear or logarithmic scales or using spline smoothing with variable smoothing constant.

3.1.1 Panel description



In the top part user can select *First set data input* and S*econd set data input*. The behavior of these controls is always the same in all Irena macros system.
Select data naming convention used – Indra 2 (Folder names for sample names and then DSM_Int, DSM_Qvec, and DSM_Error), "qrs" (suggested use of folder names for sample names [but not necessary], q_sampleName for q vector, r_sampleName for intensity and s_sampleName for error. Or do not check any checkbox and all folders and waves will be listed (allows for ANY naming convention to be used).
The pick *Data folder*. Subset of folder will be listed in specific naming conventions.
Waves with data may or may not be automatically selected, depending on naming convention.

You do not have to select error wave, but it is strongly suggested to create one during data import, if better error wave does not exist!!!

Add data and graph after selecting one or two data sets (use data set 1 as single set, please) click this button – data are recorded for use by the tool and graph is created.
Reset Modify resets the tool
Autoscale When two data sets are used and range of data (overlapping data) is selected using cursors, this will calculate area under both curves and use the ratio of these areas to scale data set 2 to data set 1. Will write the result into the *Modify data 2*, *Multiply int by*. There it can be later modified by user.

Modify data areas. Scaling factors, background subtraction, etc.
Rem Q<Csr(A) remove data with Qs smaller than Q for the data point on which is cursor A (rounded cursor) set. Note, this gets the data set name from the wave on which the cursor is positioned. Place cursor on right data set you want to modify.
Rem Csr(A) removes ONE point, on which the cursor resides at this moment.
Rem Q>Csr(B) removes points with Qs larger than Q of the point with cursor B (squared) for the data set on which the cursor B is. Note, this may not be the same data set as where the cursor A is!

Combine data joins data by simply writing two data sets into one data set and arranging them from smallest Q to largest Q. No rebinning for Qs is done. Therefore it is likely in overlap region may be more points than appropriate. Suggestions for use are below in the section titled "Various uses for this tool.".

Data1 – Data2 ; Data2 – Data1 ; Data1 + Data2 ; Data1 / Data2; (Data1 – Data2)/Data2 These do the math described here on the intensity data (and properly propagate the errors). This is done after rebinning data 2 to Q points of Data 1. This is done by interpolation of logarithm of intensity and then conversion back to intensity.

Data 1 using Qvec2 Does ONLY rebinning of Data set 1 to Q vector of data set 2 (and propagates errors, if available).

Data 1 ; Data 2 Passes data 1 or 2 only to allow use of smoothing on data without any math being done.

"Reduce No of points by" Takes Data 1 and reduces number of points by number selected (2 = ½ points left).

"Log Reduce to" Takes Data 1 and reduces number of points using approximately logarithmic reduction. Generally, at low-q more points are left while at high q more points are summed. Note, this reduction is by averaging points on log-scale and assigning them new, average, q, intensity, and error (all simple averages). This is done by first creating new log-Q scaling and then going through data and summing in each newly created q bin (borders being linear half distance between preceding and subsequent points) all of the Q, intensity and errors (if exist)… Bins with no values in them are then dropped… This has side effects: you cannot create more points then exists, you cannot increase point density (no interpolation is done) and the new Q scale may not be exactly logarithmic. Also, you are not guarantee to have the right number of points you requested, as some may have been dropped, if locally the density of new points was higher than density of old points…

Slider below the checkbox controls the weighing of the log scale – how many points are created at low and high Q values. You need to push the button "go" first time when using the "*Log reduce to*" – after that the slider updates the graph every time the value is changed and mouse is released.

This tool can be very useful, when large number of points exists at high Q with very little information content. Especially when input data have very high number of points, some of the tools will run very slow or not run at all for memory limitations. This data reduction may be the right choice…

Results of above mathematical functions can be also smoothed by checking *Smooth* checkboxes and selecting appropriate *Smoothing window*.
Smooth (log) smooth logarithm of intensity and then converts back to intensity
Smooth (lin) smooth directly intensity
Smooth window number of points accounted in "running average" smoothing used here…
Smooth spline uses spline interpolation (with slider on right setting the spline smooth parameter). Use carefully – and test right setting of the slider. You need to push "go" first time, after that (when smoothing parameter is changed and mouse is let go) the data are redrawn automatically.

Output area
Pick new data folder pull down menu allows selection of existing folder in Igor experiment
New data folder modify, write in data folder name (fullname, staring from root:…., use the above pull down to preselect). If the folder does not exist, it will be created
New intensity name, New Q wave name, New Error wave name names for output waves. Please fill in appropriately for your naming convention**.**

3.1.2 Various uses of this tool

1. Merging segments from various sources. Import segments and then load in as Data set 1 the set which is calibrated, as set 2 next one and using automated method (select data overlap by cursors) scale them together. Remove extra points, reduce number of points in overlap area (if necessary) and subtract backgrounds, if necessary. The *Combine data*. Smooth if desired.
2. Subtraction of one data set from another. Load the data sets in, scale and modify if necessary. Do the proper math. Smooth if necessary.
3. Smoothing data. Select only data set 1 data, pas it through math part (modify, if necessary) and select smoothing method desired.

Please, make sure you fill in properly the names of waves for output data. There is nothing I can do here to catch typos and mistakes…

3.2 Data manipulation II – many data sets

This tool serves for modification of many data sets at once. In first version it can ONLY average waves, but it can be modified to do other things. If you have any ideas, let me know.

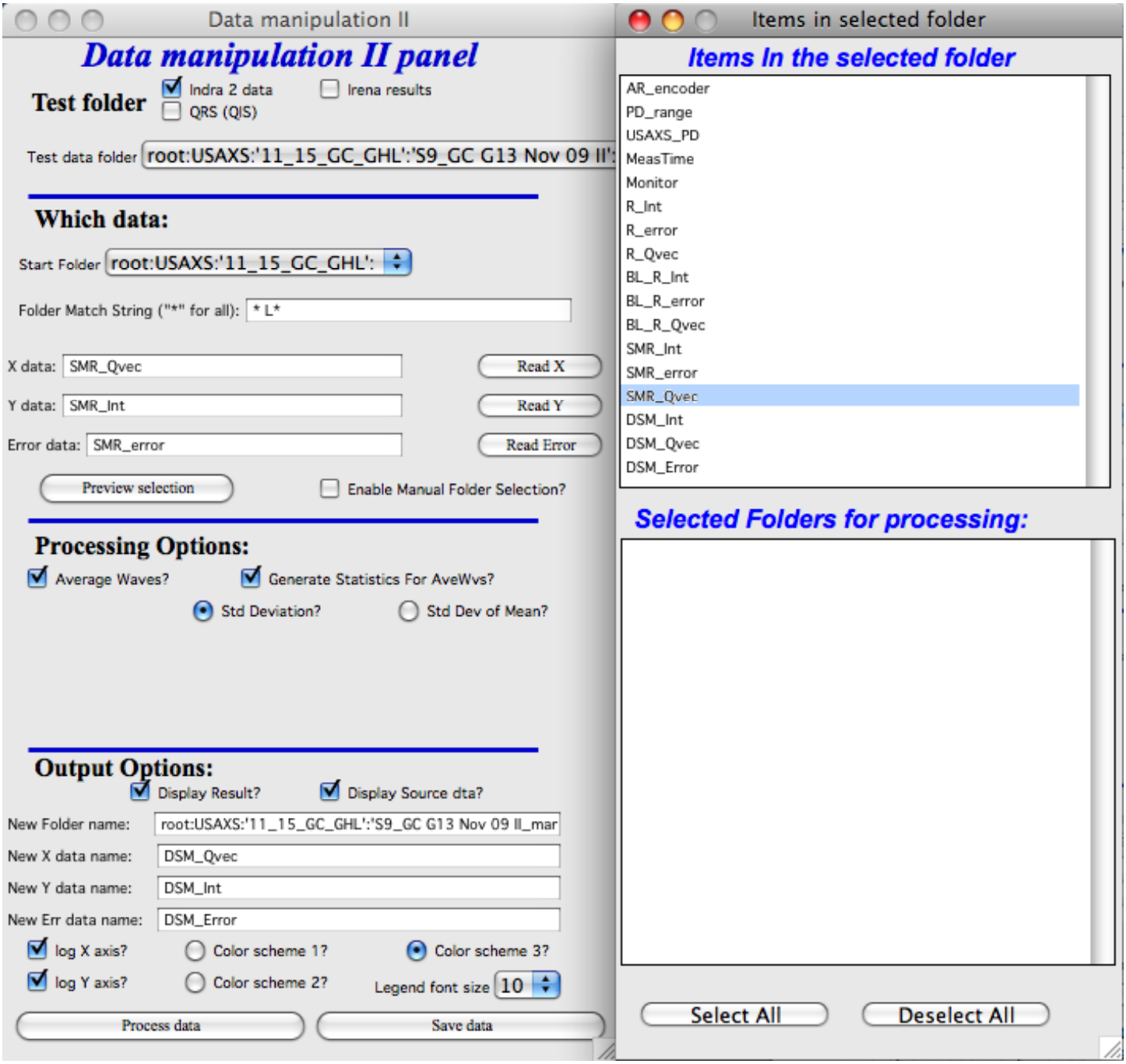
3.2.1 Introduction

This tool is combination of data selection method used in Data Mining tool, which enables user to select folders with data. These data can be processed (averaged for now), plotted and saved.

It is relatively simple tool at this moment… User selects by one of few methods number of data sets and these then are processed.
Averaging – first data set found during processing X (Q for SAS data) values are used for result X vales. At this time there is not way of reordering the data, so this is kind of random which X wave is picked. Then for all data sets the Y is linearly interpolated for each X value and average and standard deviation (or standard error of mean) are calculated. It is not necessary to have the same number of points – and it is even not necessary to have same X (Q) range, but if the X range is different, Igor will extrapolate the closes existing values as fixed number. That is likely going to make such situation really bad. But anyway, the code will not fail.

What is nto done yet: As written now, the code simply dumps any errors provided and uses only the standard deviation on the Y values. If I get smarter, I can try to propagate the errors also.

3.2.2 GUI and controls



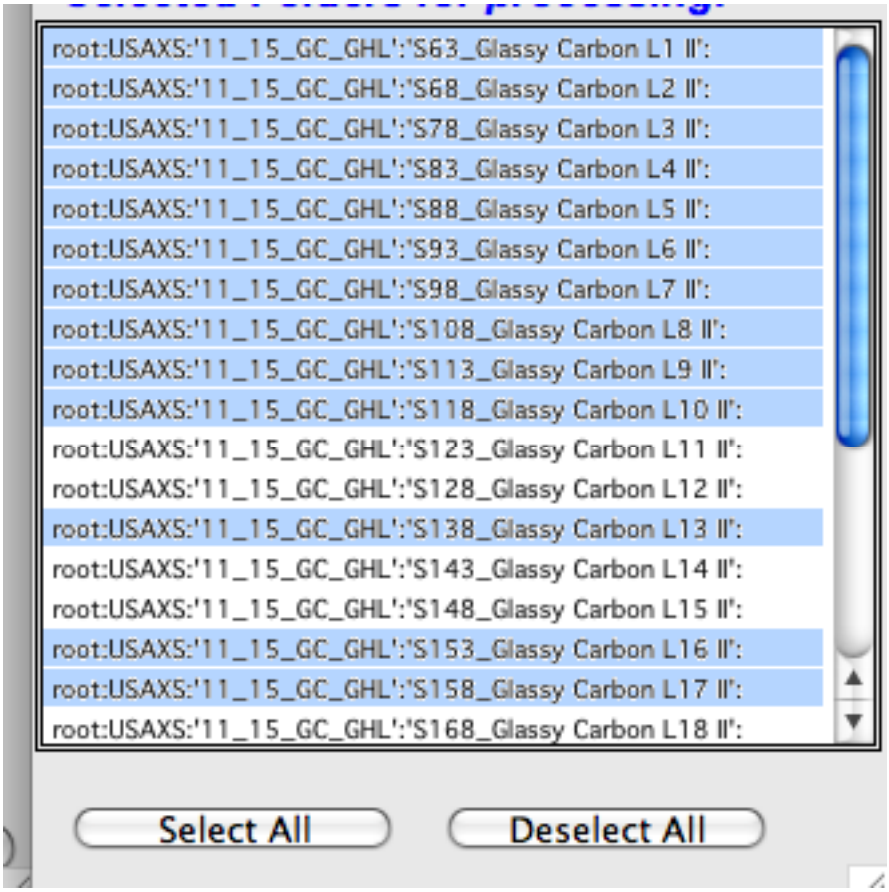
The control panel is on right, Panel with two list boxes for user controls on the left.
At the top of the control panel is standard suite of data selection control. User should select folder in which are waves which user may want to do something with. In my case I selected a folder containing some Indra (USAXS) data I want to sum together.
Please note, that for lack of better choice, this "Test Folder" name will be used as template for output suggestion – basically, I will add "_manII" to its name. You can modify later.

Which data
In this area user should try to select as best as possible the data to be used.
Start folder – Lists folders in the current experiment, any folder below this one in the folder hierarchy will be searched. Therefore, user can select only part of the data to be searched and considered.

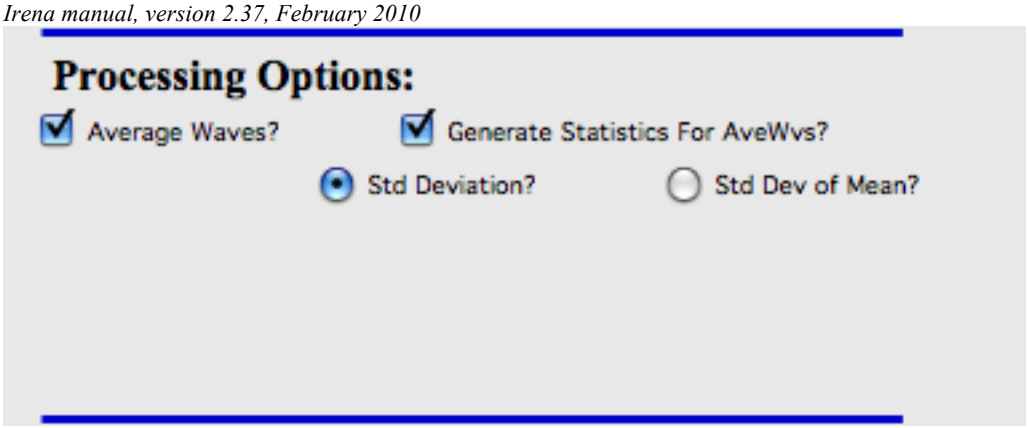
Folder Match String – string which can be used to match to folder names. Uses more or less standard system, basically I suggest considering this: * matches any folder name, *K* matches any folder containing k in the name (the search is not case sensitive. In the example, I wanted to match all folders which contained space and l next to each other, so I used * L*. Note, do not add "" to this string.
X data, Y data, Error data – you can use either full name (like the Indra data "SMR_Int", etc.) or you can use * to create string which matches part of the name. For QRS data, for example, you can use q*, r*, and s*. While this does nto guarantee the right triplet is used (firs wave starting q will be used for x, first starting r for y etc.), it should usually work. Things get messy if many qrs waves are in the same folder. This tool just cannot manage that. If you have weird system, send me example and I can try to fix it.

Preview selection – runs data checking code and finds the folders with data. Just the folders. They are listed in the bottom listbox in the panel on right.

Enable Manual selection – if selected, user can manually select only subset of folders in the listbox on right (bottom). Use buttons *Select All* and *Deselect All* - as well ctrl and shift – left mouse button to select ranges of data.

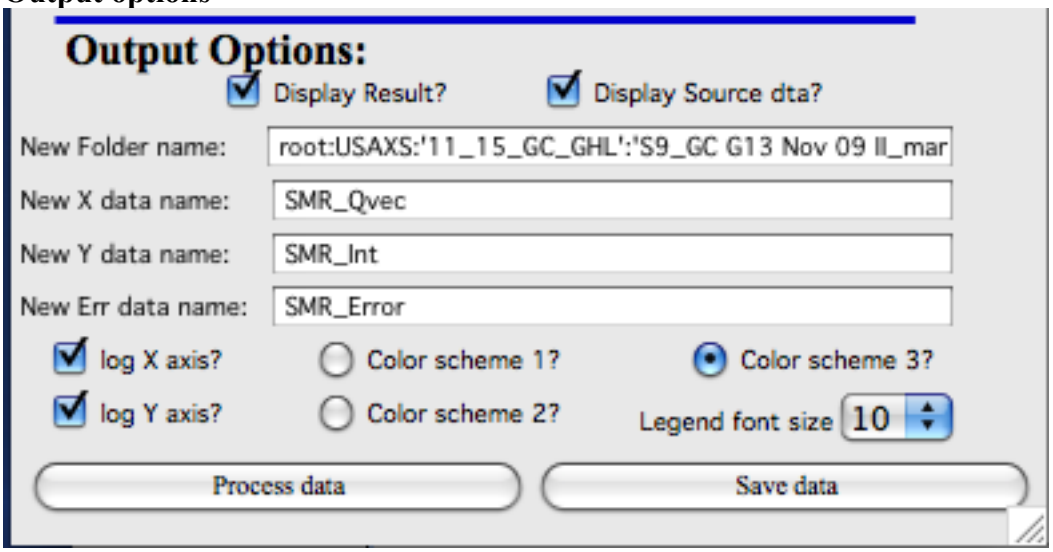


Now the Processing controls:

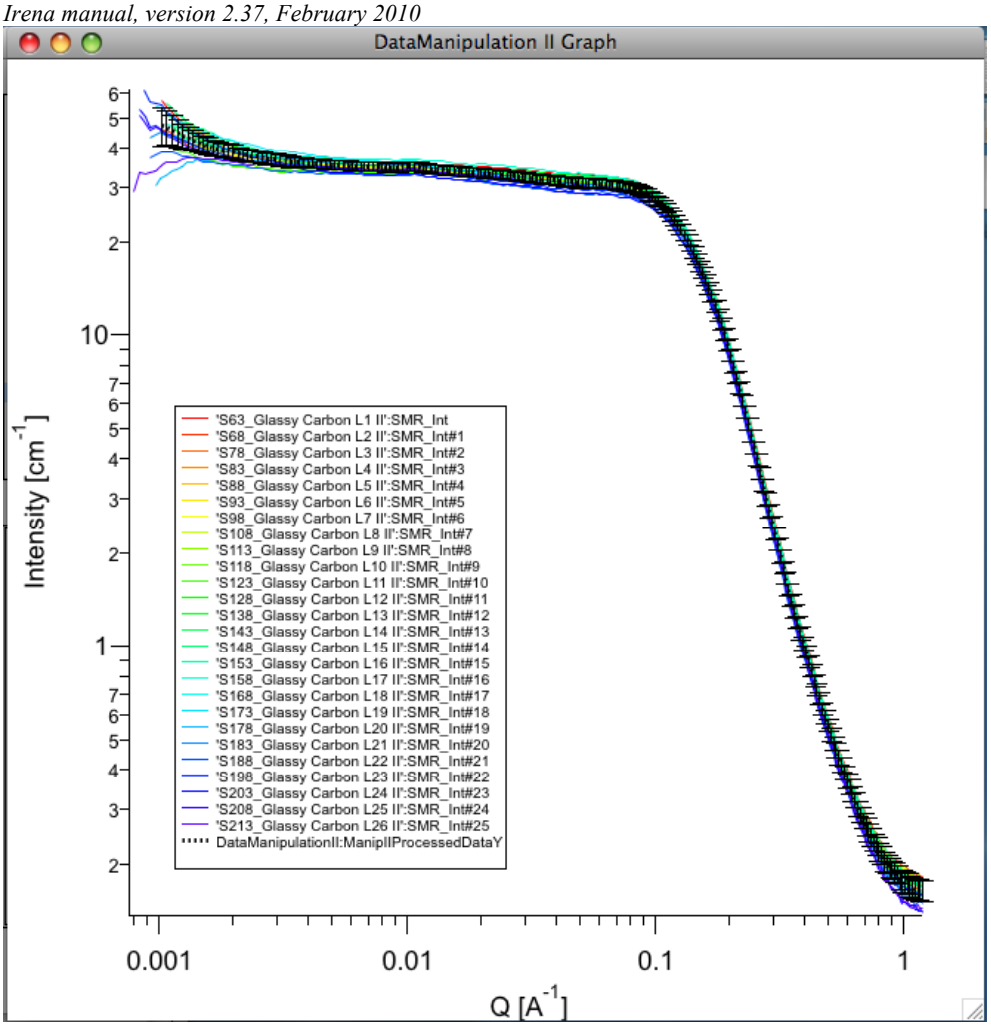


Hopefully later there will be more options. For now you can *Average Waves*, and you can choose to create statistics (either standard deviation for each point or standard deviation of mean). As noted above, at this moment this is purely statistics on Y values, Errors are not considered.

Output options



Display result? **And "*Display Source data?*** Will cause that a graph with results and source data will be presented when *Process data is pushed.*



Example of plot with data with average.

New Folder name and X, Y, Err names - folder needs to be with path (keep it short), separated by :. No need to add ‘ ‘ to names with spaces, the code will fix it. If a name is too long (more than 30 characters) it will be cut short. Wave names are simple strings, can contain spaces, but no special characters. No +, -, and other weird symbols.
Other controls below control how the output graph looks like.

If the results look good and you like them, use *Save data* button, which will store the data in the folder and under names in the above controls.

4. Plotting tools

4.1 Introduction

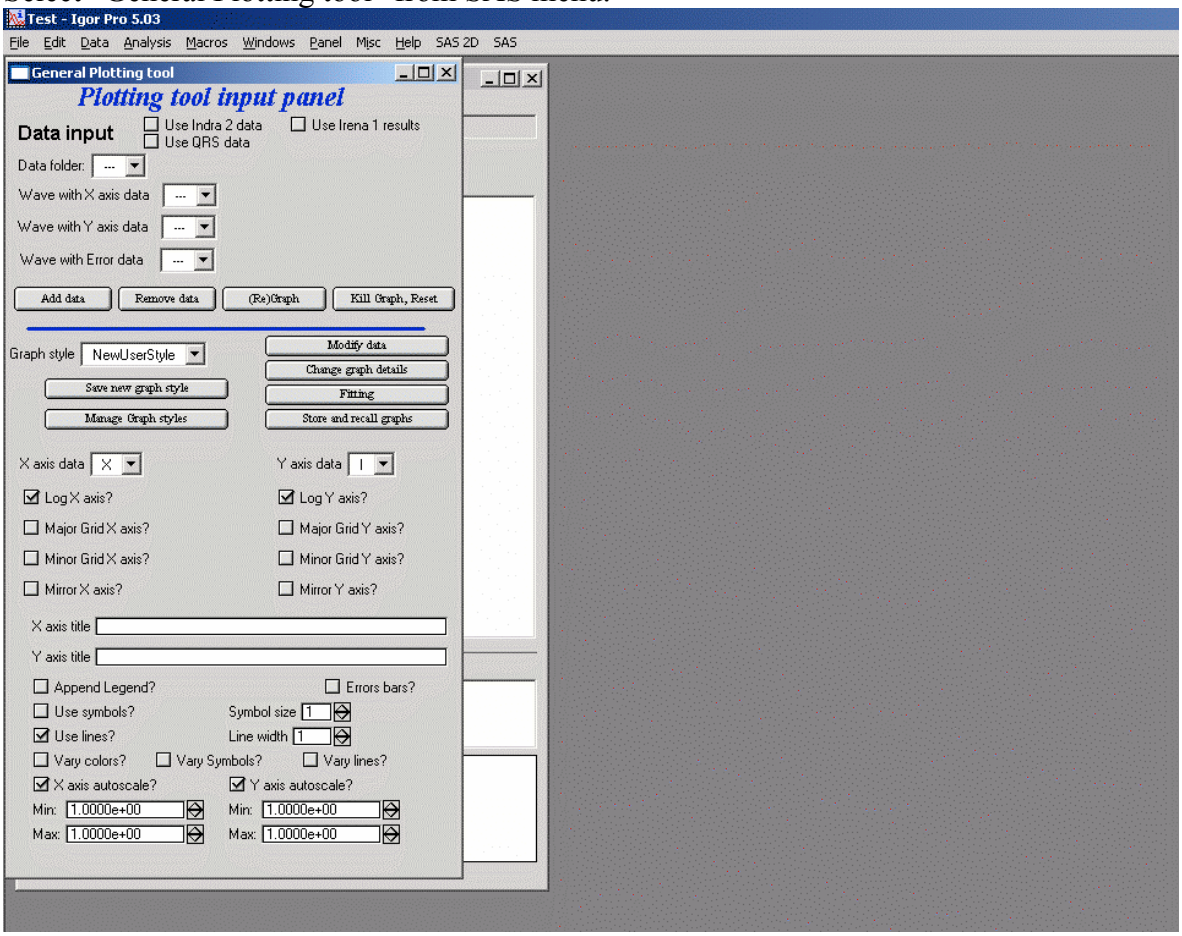
This plotting tool is designed to be able to produce publication ready plots of SAS data and simple fits. The tool allows creation of plot user styles, which can be applied quickly and reproducibly to numerous sets of data. New data types can be automatically created. Please note, that the formatting is saved ONLY if it is done through my custom made panels designed for this purpose.
The tasks, which can be done:

1. Load data and plot them, new data types (e.g., Y * X^4) are automatically created if necessary.
2. Modify data (Multiply Intensity, remove points, subtract background,…)
3. Do simple fitting (Porod, Guinier,…)
4. Create plot user styles, import and export them from current Igor experiment for future use

Note, that this tool also allows displaying "results" – size distributions, Unified fits etc. The capabilities are still little bit limited in this area… I plan to improve on this.

4.2 Description

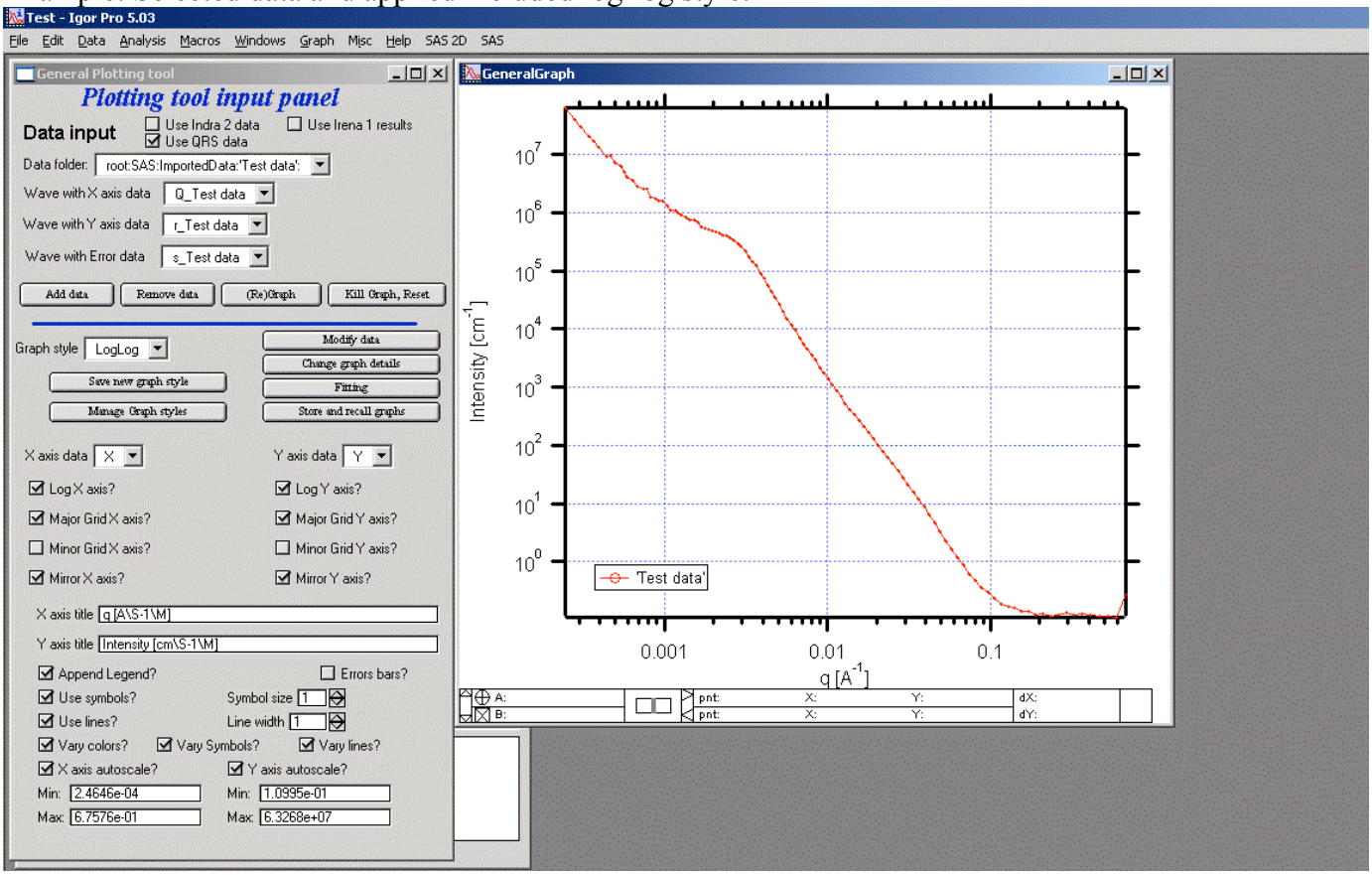
Select "General Plotting tool" from SAS menu.



Select data in usual way and push button "Add data". You can add multiple data sets (most formatting in this tool is set for up to 8 different data sets). You cannot have the same data set twice, code will complain and refuse to do so. You can add data anytime later also.

Apply Graph style available in the popup "Graph style" or select the data type to plot on both axes. The needed data are created, if they do not exist.
Use checkboxes and more controls in "Change graph details" (opens new panel with more space) to modify graph as needed.

Example: Selected data and applied included log-log style:



Note few items: In the axes names you need to use Igor formatting for subscripts, superscripts, Greek letters etc.

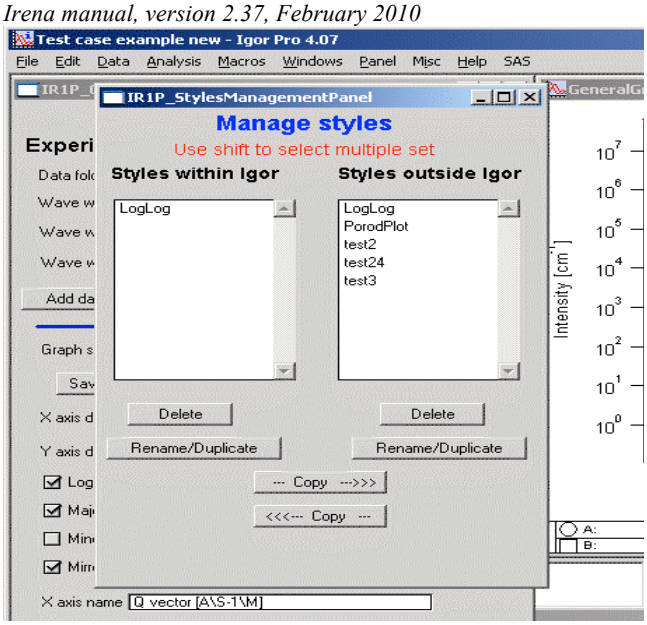
Setting limits on the axis can be done manually or using function which can be called using zoom function in Igor. Select area of the graph you want to zoom to and right-click. Select "ZoomAndSetLimits" from the menu. This will zoom the graph as well as set limits in this tool.

Creating user style

When you have graph which you like to use many times, click button "Save new graph style". The new style will be created after user provides name. The name is checked for uniqueness and for name appropriateness, so the new name may be slight modification of the name provided. You can rename the style using "Manage Graph details" button.

Import & Export of styles

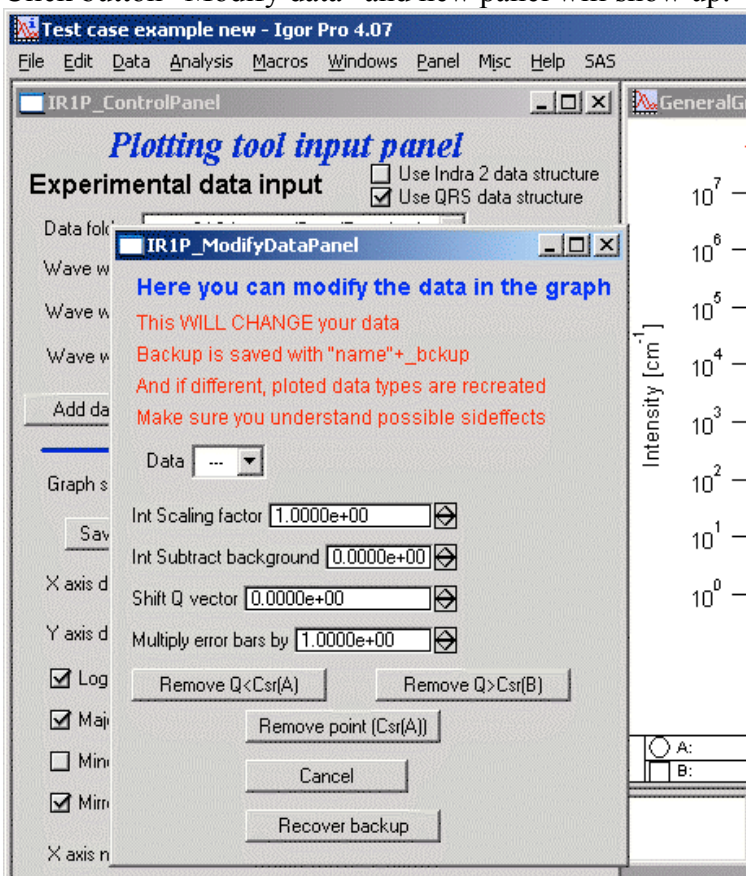
Use "Manage Graph details" button.



The panel shows two main lists. Left shows user styles available in Igor and right shows styles outside Igor. Buttons under each window allow manipulation with the styles, the buttons " --- Copy -→" or "←-Copy ---" can be used to copy styles between the Igor experiment and hard drive storage space.
When done, kill the panel.

Modifying the data

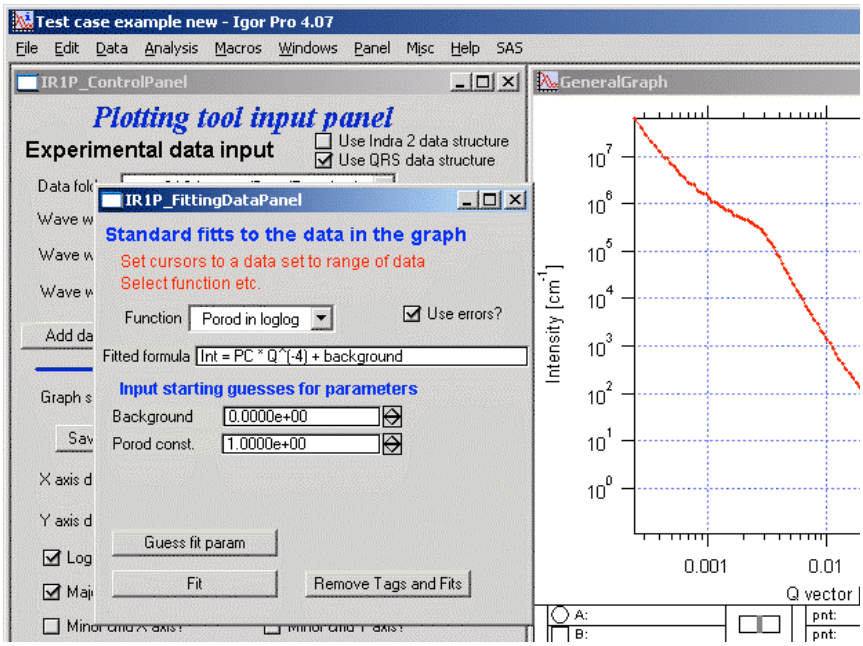
Click button "Modify data" and new panel will show up.



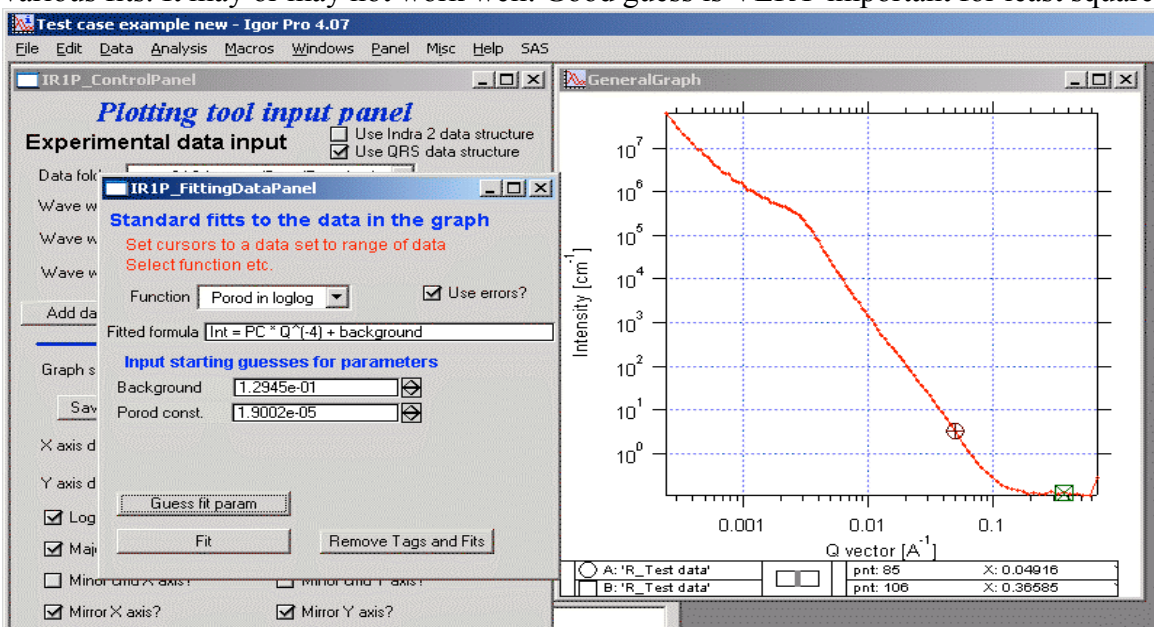
Important information:

When this tool is used **FIRST** time on a data set, it creates a backup copy of the data. Anytime later, this can be recovered. If that is done, **ALL** changes done to the data will be removed.
Select data to modify, modify using buttons and numbers. For removing data smaller than particular Q or removing just one data point, user rounded cursor (called A in Igor) and for removing data larger than particular Q use cursor B (squared).
Use button Cancel to reset the corrections to default values seen above. Note, that this resets instance of running this tool, to reset data to original data you may have to recover backup of the data…

Fitting

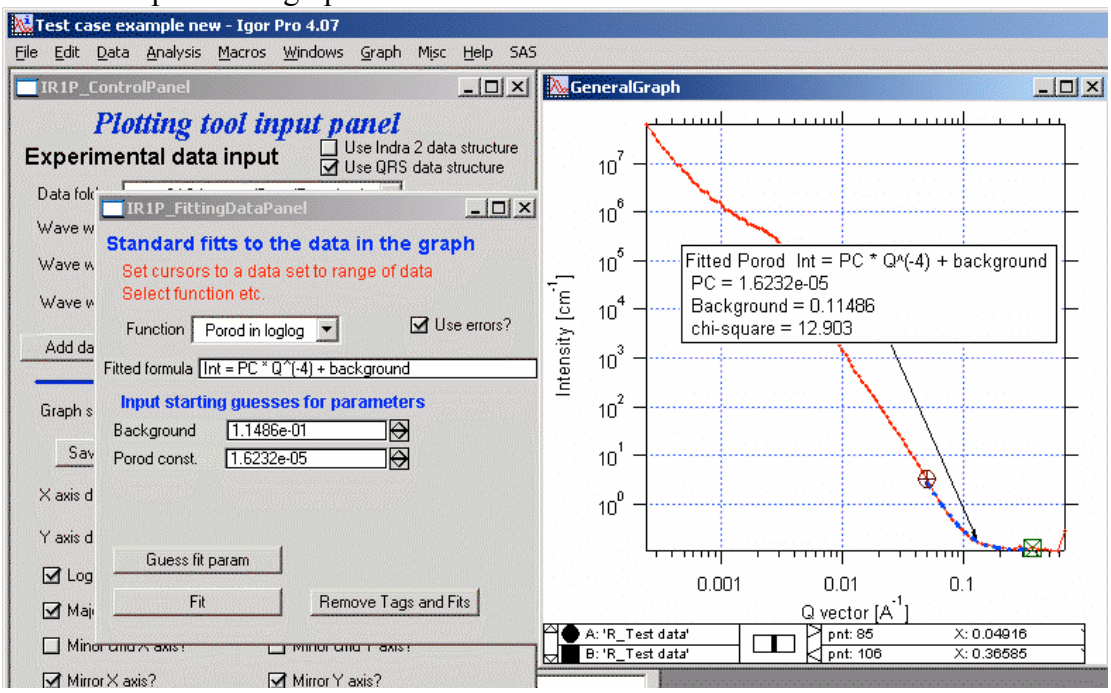


Use button "Fitting" to pull up panel with fitting range, checkbox "Use errors" if these are available. Select function to fit. Use cursors to select range of data to fit.
Note, that various number of parameters appears below for starting guesses for parameters. You can try to push button "Guess fit parameters". This will try to guess starting parameters for fitting, since these are necessary for various fits. It may or may not work well. Good guess is VERY important for least square fitting. Example:



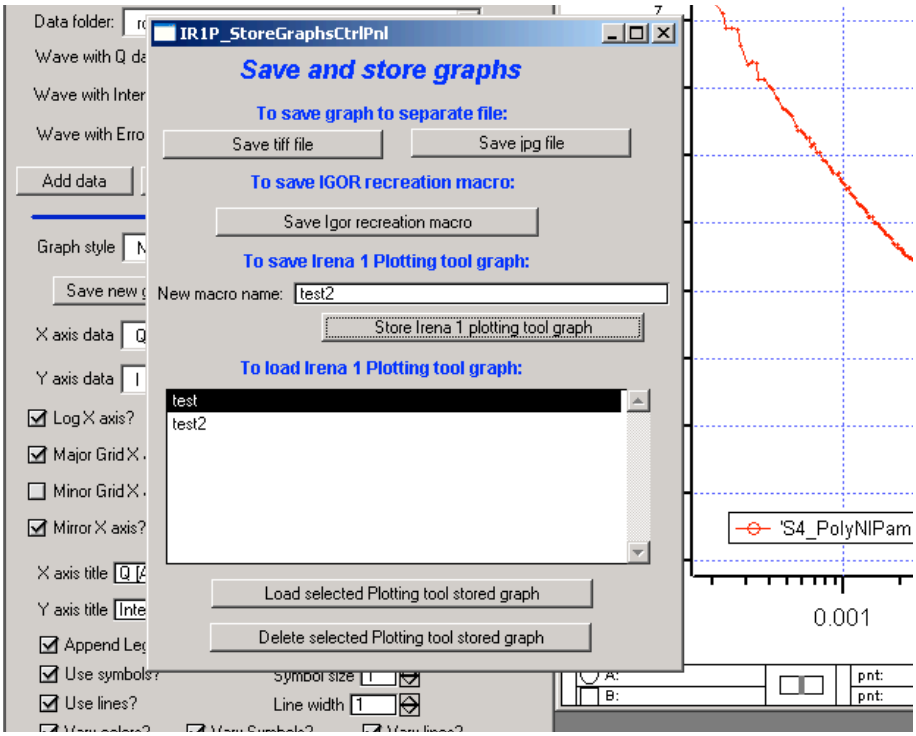
The try to fitting:

Results are printed in graph:



To remove the results from the graph and clean it up, use button “Remove Tags and Fits”.

Storing graphs for future use, exporting figures



To get a control panel controlling the features for storing graph, exporting graph and recalling stored graph, push button on main screen “Store and recall graph”. This pulls up the above control graph.

Controls description:

The two top buttons allow user to save current graph as tiff or jpg files. The dialog for naming them is provided after pushing the button.

To save Igor recreation macro, push the next button. It does not work yet… I need to figure out how to do this.

Next is name for Irena own recreation macro, which you can store. This macro is in form of string and stored in "root:Packages:StoredGraphs:".
The advantage of using this macro compared to Igor Pro recreation macros is the fact, that after recovery of graph through this macro the Plotting tool can still control all features. That is not true for Igor recreation macros.

"Store Irena plotting tool graph" button will store the current graph in the above-mentioned place as a string. The strings are listed in the listBox below.

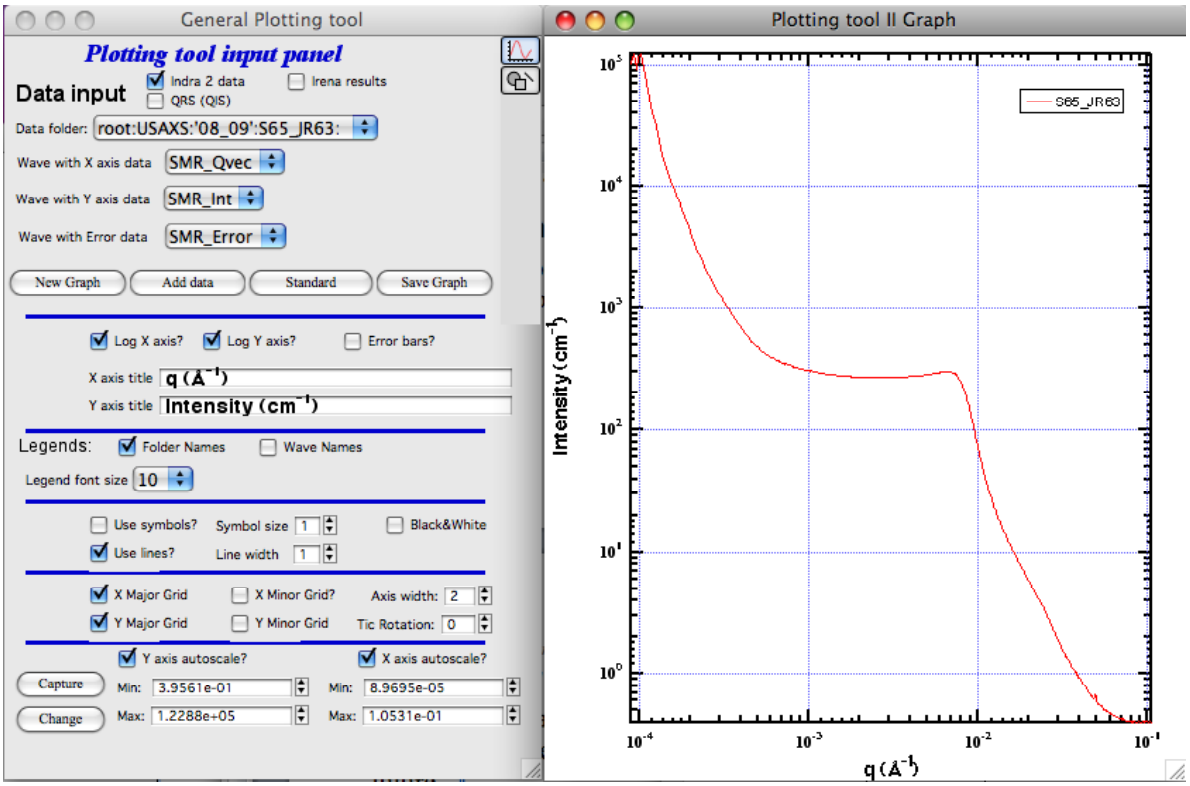
Selected stored graph in the listBox can be either restored – or deleted, using the buttons below.

Modifying the data in the graphing tool

To restart the tool and cleanup the data from graph push button "Kill graph, reset", to remove one data set at a time from the tool, use button "Remove data".

4.3 Plotting tool II

This is modification of plotting tool developed by Dale Schaefer. This tool control ANY top graph. It can, therefore, be used more flexibly – but has some limitations… This tool is likely to be developed more in the future.



The GUI contains selected controls and any change in this GUI is applied to the top graph. Note, that compared to Plotting tool I, which at each modification reapplies all formatting to the graph it controls, this tool applies only the control which has been changed. Therefore, this tool is basically ONLY different GUI to Igor controls – combined with convenient Irena data selection tools.

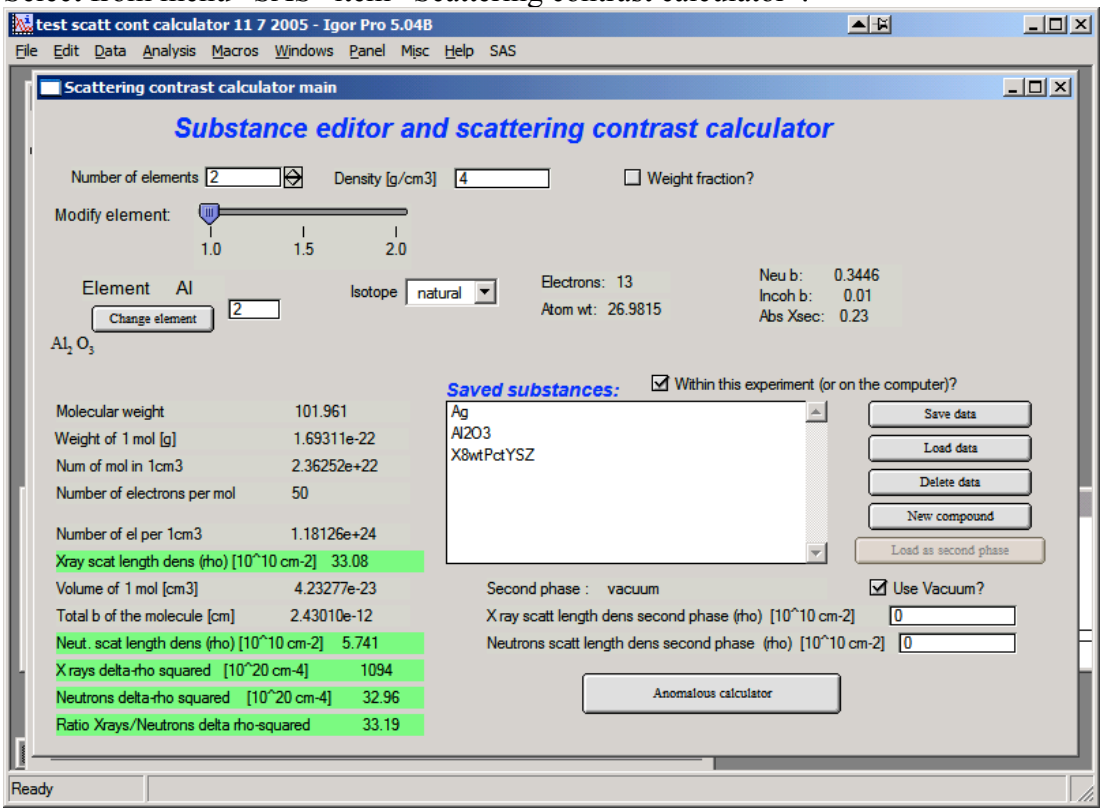
5. Scattering contrast calculator

5.1 Introduction

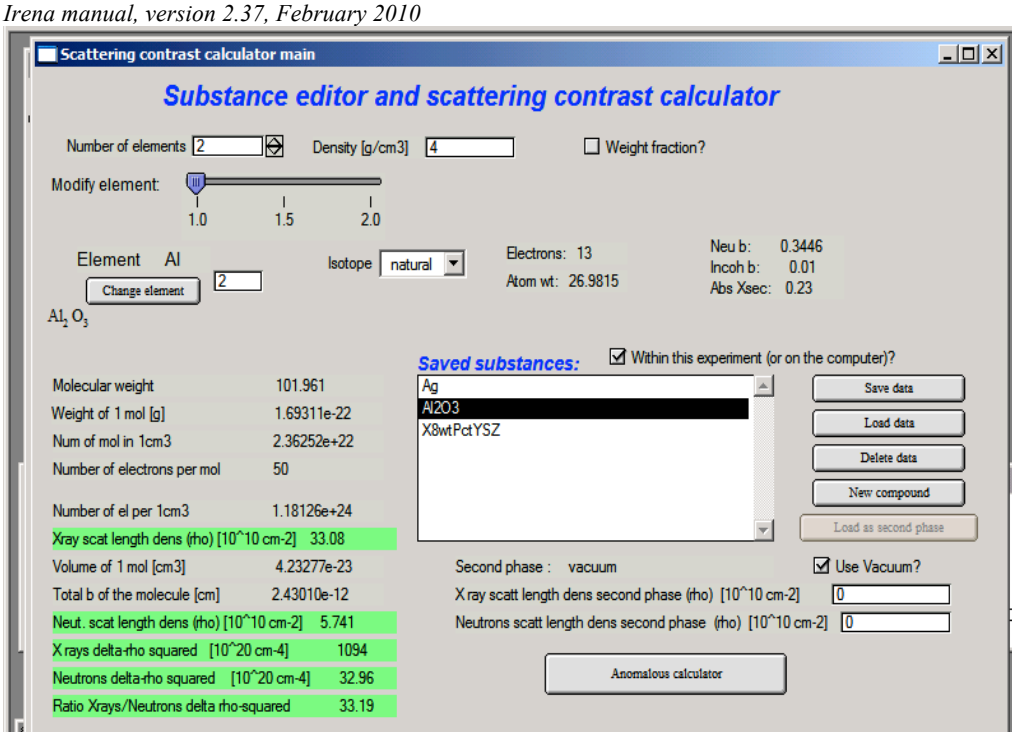
Calculating scattering contrast for various compounds is annoying “monkey” work, which can be easily left to computers. The tool in “Irena” at this time allows one to calculate the X-ray and neutron scattering contrast for compounds with up to 24 atoms, with known density, with known atomic fractions and **with no energy dependence**. Energy-dependent anomalous scattering effects (for X-rays) may also be considered by an additional level of this tool (see button marked "Anomalous calculator" in lower right corner).

5.2 Running Scattering contrast calculator

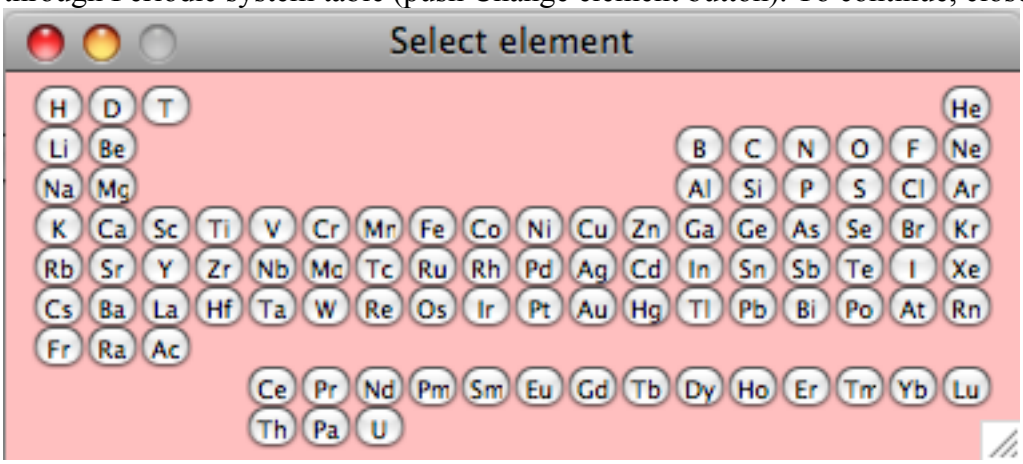
Select from menu “SAS” item “Scattering contrast calculator”:



This is the interface. At the top, select number of atoms in the material, set its density and check the check box if you want neutron data displayed. Let’s select 2 atoms, may be Al_2O_3 (Corundum) with density of 4 and see neutron results. This is the picture after this selection



Use slider to select each element and check it's properties – amount in molecule, Isotope etc. Input is done through Periodic system table (push Change element button). To continue, close the table…



Most of the fields is filled automatically with data from internal databases of this tool. In the lowest part of the tool are results and intermediate calculations of this tool – so one can obtain various numbers, which needed to be calculated.

5.3 Use of matrix

To calculate delta-rho squared … $(\rho_{matrix}-\rho_{scatterer})^2$ ... we need to set scattering length density of matrix. This can be done in numerous ways:

1. Write the numbers directly in the fields provided
2. Calculate the matrix scattering length densities and use "Set as matrix" button
3. Save matrix data using "Save data" button and then load them as matrix "Load matrix data"

In each case the values for "delta-rho squared" should be recalculated. Note, that if checkbox "Use vacuum as matrix" is checked, vacuum is used as matrix and no selection for matrix is available…

5.4 Saving data

This tool has "saving" feature, which allows to save the compound parameters in such way, that it can be used in the future. The data **can be saved either INSIDE the current Igor experiment or OUTSIDE Igor experiment**. Compounds saved outside ***are available to any Igor experiment on that particular computer***. But experiment moved to another computer will not have these compounds saved…
To select where compounds are saved, use checkbox "Within this experiment(or on the computer)?
Use buttons "Save data" to save current compound, modify name as necessary – keep in mind to keep the "" around the name and use characters allowed as file names. Limit name to 27 characters or so (Mac name limitation).
Use buttons "Load data" to load data in the tool and "Load matrix data" to load data as matrix ONLY…

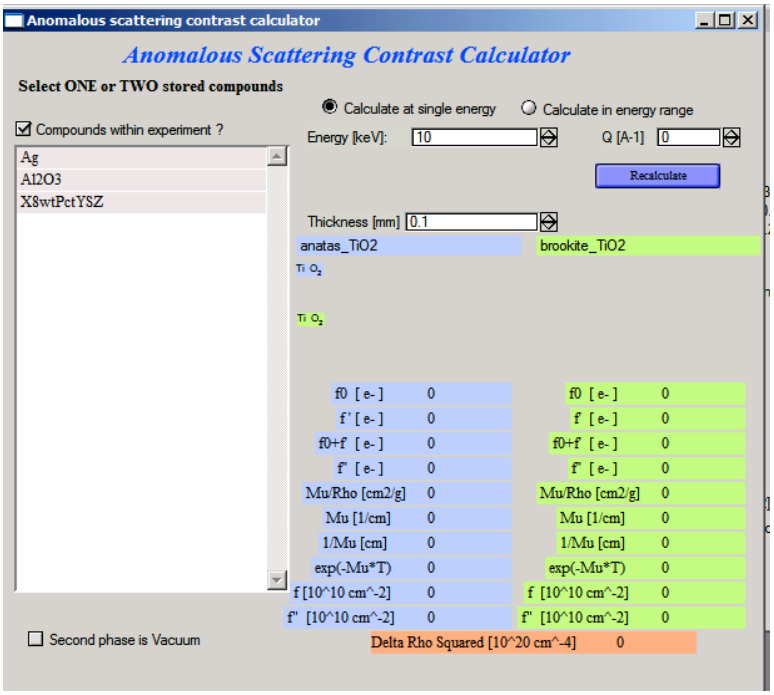
Comment: Due to rounding related to saving the data in ASCII file, there will be rounding error when using "Load matrix data" in the "delta-rho squared" calculations…

Button "New compound" will clear all settings in the tool to start creation of the new compound.

New comments on saving the data: From this release the compound data are saved with in the same place where the Irena macros are stored. This is to allow users of limited privileges to run and operate. See above comments on macros installation.

5.5 Anomalous calculator

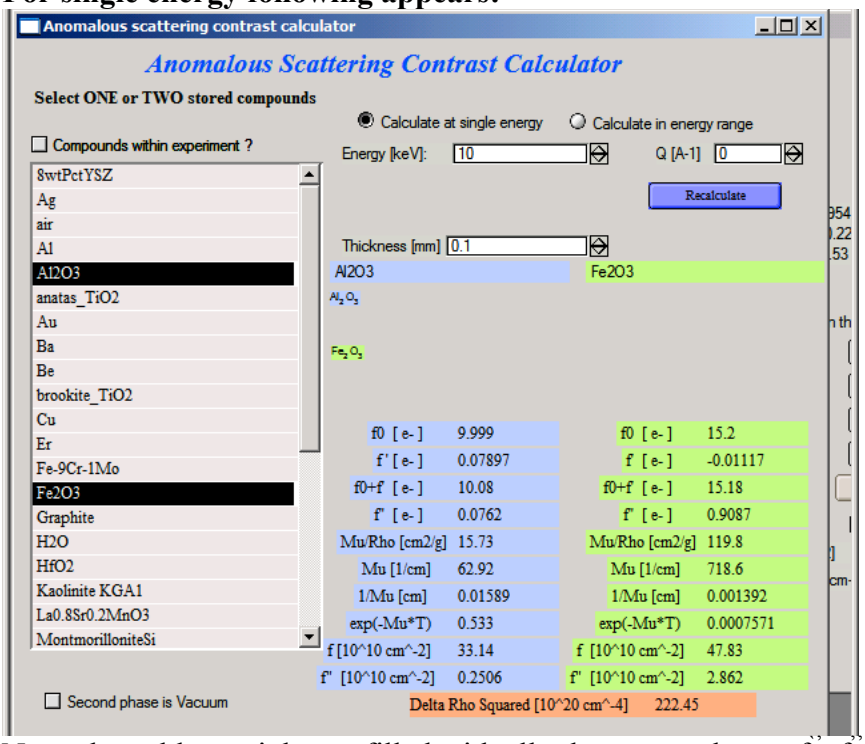
The package includes Cromer-Liberman code for calculating energy-dependent (anomaouls) effects. The button "Anomalous calculator" on the "Substance editor and Scattering contrast calculator" calls up new window…



Use of this tool:
Select one or two compounds created and SAVED in previous (regular scattering contrast calculator). If you select only one, use vacuum as second phase (checkbox below the selection of compounds). Then select, if you want to calculate values at one energy or in energy range. Note, that calculating values for large number of points may take quite a long time.

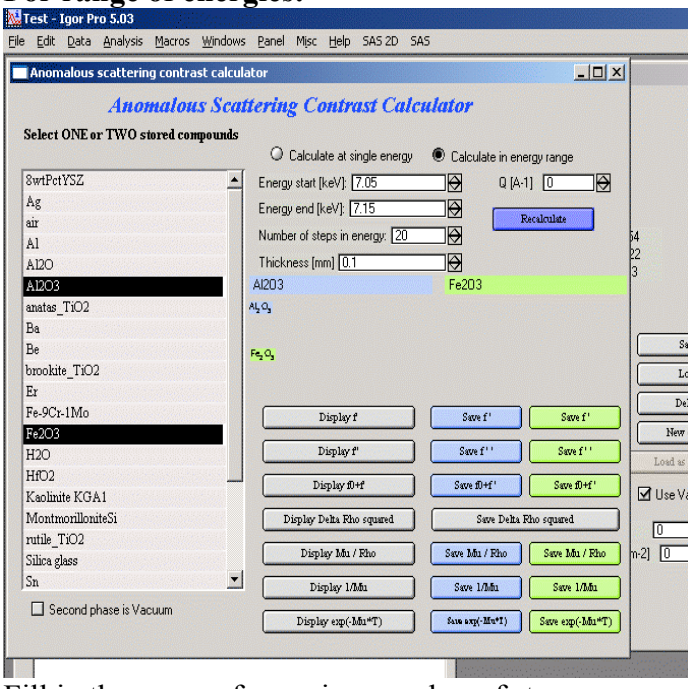
To select two compounds hold shift. Then input right thickness and click "Recalculate". Fill in the Q if you need values at higher Q values (for small-angle scattering assume Q=0)…

For single energy following appears:

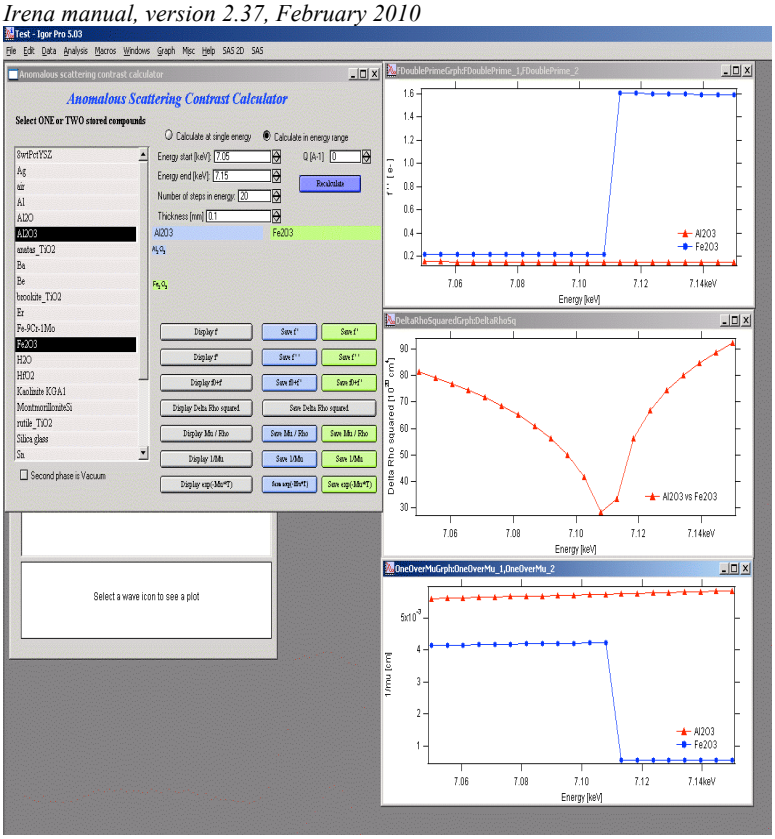


Note, that table on right got filled with all relevant numbers – f'', f', μ, and related values for each compound separately. Note, that f' and f'' are here with two different units as output – in electrons per molecule unit and in 10^{10} cm^{-2}. Lowest number is delta-rho squared between the two compounds at this energy…

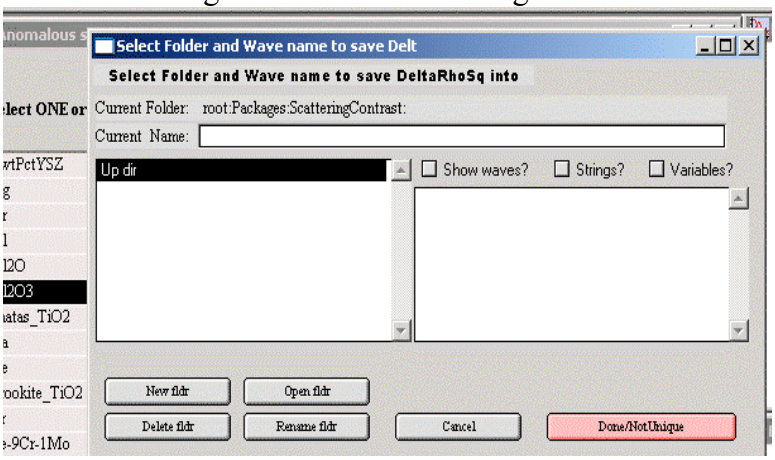
For range of energies:



Fill in the range of energies, number of steps you want to calculate (equidistantly spaced between min and max energies) and other parameters. The push "Recalculate".
The buttons "Display" create graphs of appropriate parameter, see for example below:



Buttons “Save …” save the wave with the data into Igor folder of users choice. The dialog should be easy to use and allows user to create new folder, select name for new data etc. Note, the data are saved as waves with so called “x-scaling”. To learn more read Igor manual.

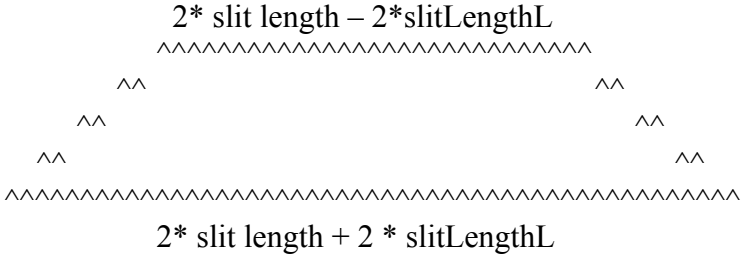


6. Desmearing

Desmearing routine built in this package is using Lake method (reference), which has been originally programmed by Pete Jemian and then coded in Igor by me. There were some minor improvements over the years, but generally this method has proven itself many times to be robust and reliable. We have verified the function of this method by collecting data from the same sample using both slit-smeared and 2D collimated USAXS. We have verified this method repeatedly and every time the desmearing was blamed for artifacts and unexpected results, we have found another reason for problems. That said, desmearing is always going to increase noise on the data… Note, that the routine will correctly handle data with absolute intensity calibration.

3/10/09 Change in desmearing tool. Per request the tool now allows both slit length (in direction perpendicular to the q direction) and slit width (in direction parallel with q direction). Further, the slit can now have shape of trapezoid, similar to what GNOM allows for instrument geometry. PLEASE NOTE: for historical reasons the parameters for Irena desmearing are ½ of the GNOM parameters.

This is the graph:



same geometry applies for slit width.

Once more, if you have parameters used for GNOM, you have to divide the numbers by ½.

The GUI changes should be easy. Please note, that :

1. If you set slit length or slit width to 0, you assume infinitely high resolution in that direction.
2. If you set "L" parameter to 0, you assume the shape is rectangular in that direction, not trapezoidal.

6.1 Theory behind the Desmearing Procedure

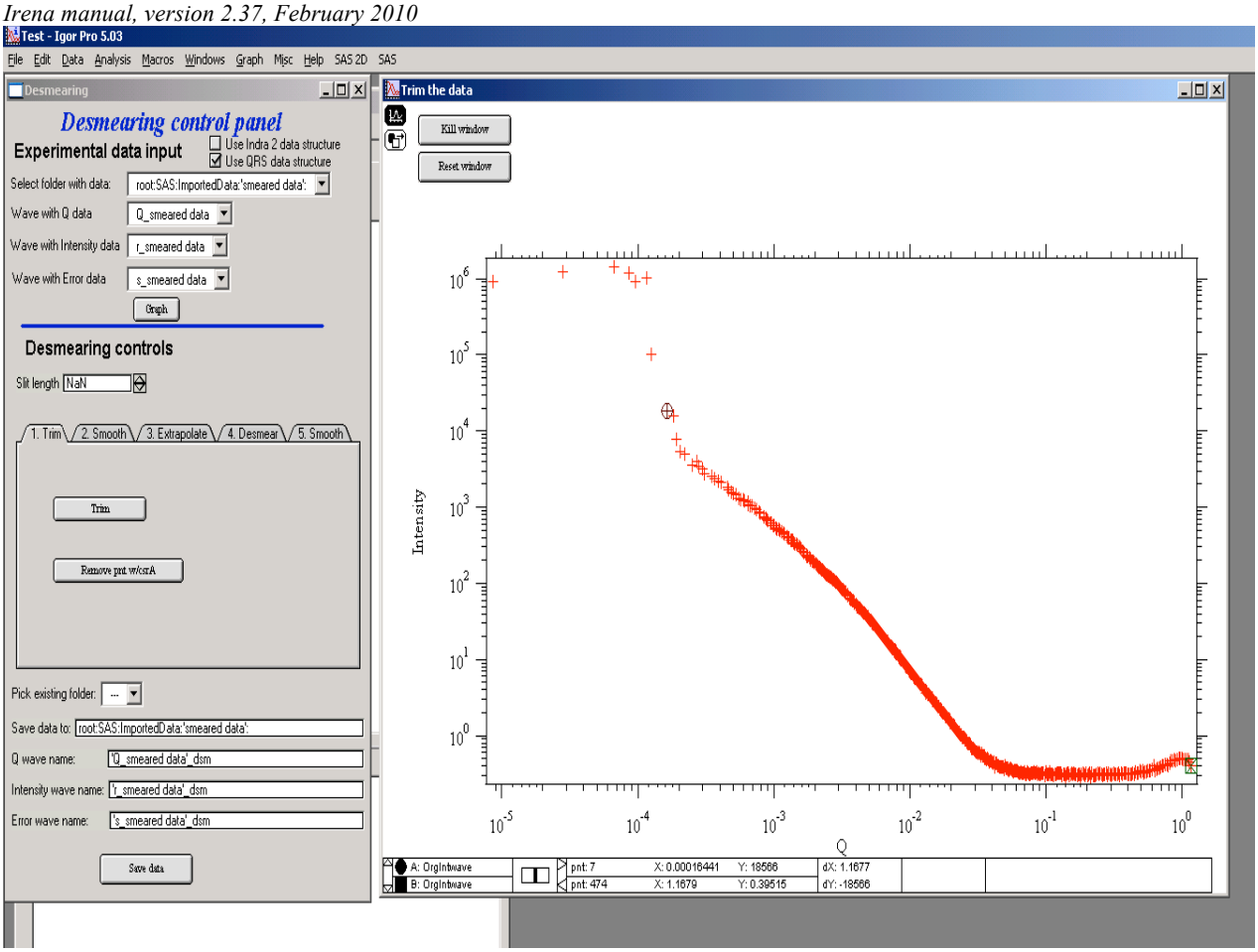
See the Lake paper.

6.2 Example of the Desmearing Procedure

I have included a file with an example data set with slit smeared data (smeared data.dat) where the slit length SlitLength=0.05113. You can include this in your experiment using Data import tool…

6.3 Final comment

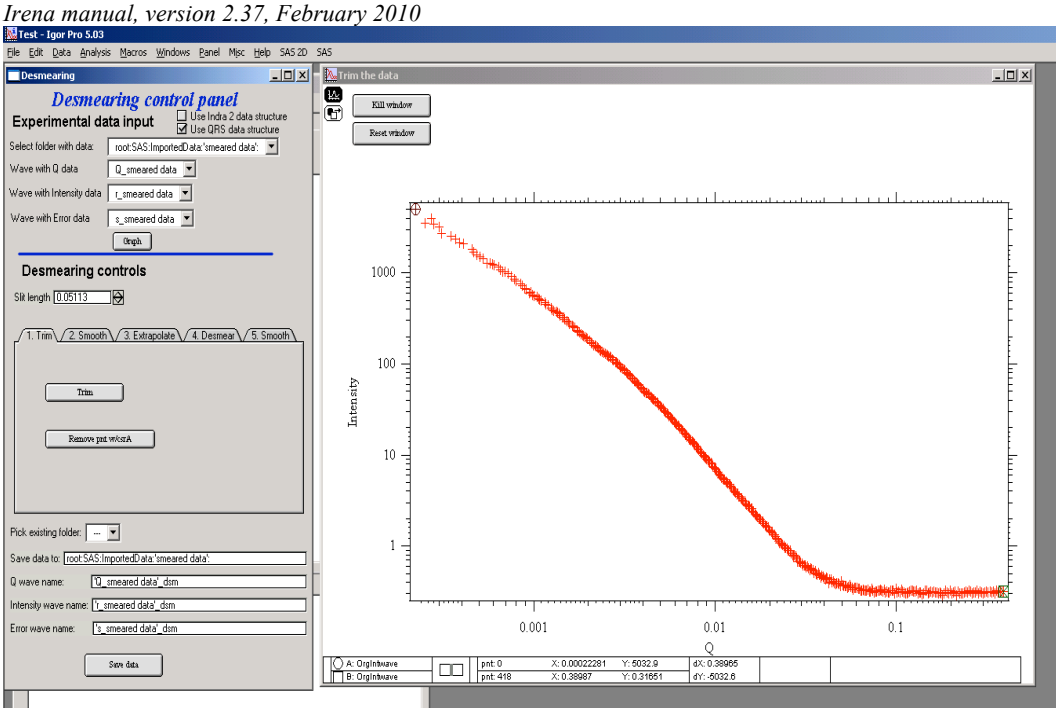
One major comment – if you need to go back in the routine, anytime you can click on previous tab and return to that place… All from tabs to the right is forgotten and routine restarts on the tab, where you click. It is also possible to skip the smoothing tabs without any penalty – note, that if the smoothing parameters are set (the checkboxes are checked) the data WILL BE smoothed, even when you do not click on the tab…



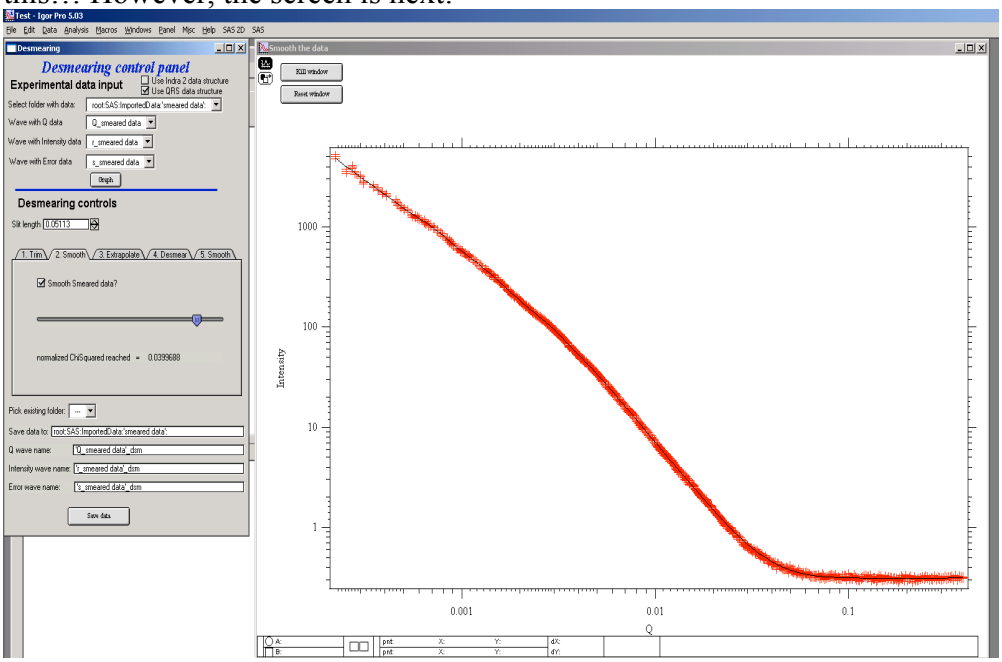
This is GUI and graph after loading data. Only thing needed is to fill in the slit length.

The tool is controlled by the tabs. The order which needs to be followed is the tabs from left to right. For each data set to be desmeared, this procedure must be followed, selecting in sequence the tabs from left to right.

1. First step – trim useable data – small and high Q data… Use cursors to select data range. And then push button "Trim". You can also remove any spurious point with the other button and cursor A (the rounded one)

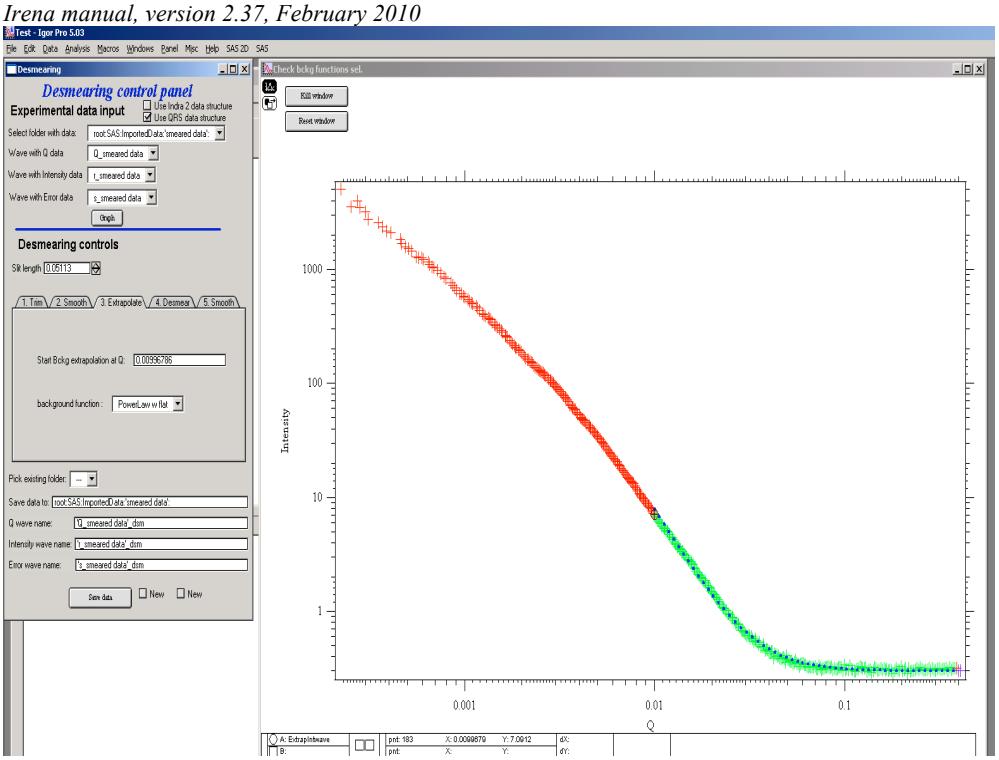


2. Next step – it is possible to smooth data using spline smoothing, but only if necessary. I strongly discourage this… However, the screen is next:



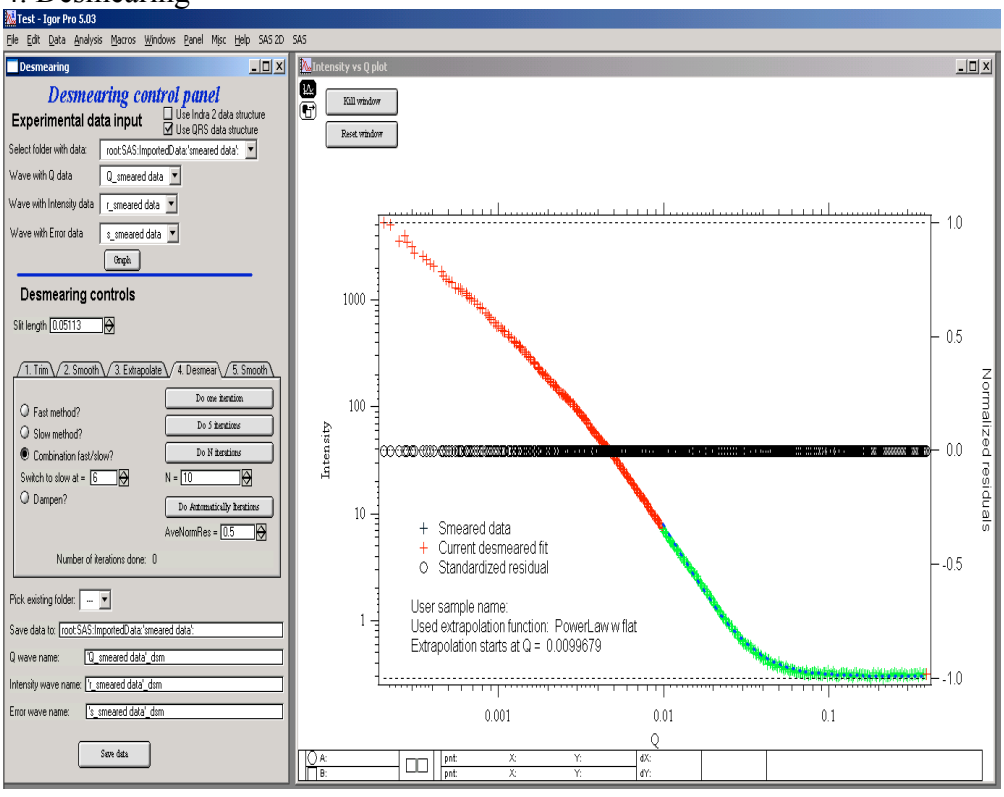
Note the slider and checkbox – the checkbox switches on the smoothing, in that case the slider appears. The slider controls the internal smoothing parameter - more to the right, more smoothing…
As I said, I discourage this, so let's remove this in next step.

3. Extrapolating.
In order to properly desmear, I need to smear and that means I need data for at least 1 slit length BEYOND the last point. Therefore we need to extrapolate the data using one of selection of mathematical functions. Most useable one is "Power law with flat" and "powerlaw" or "flat". These data suits best the Powerlaw with flat…



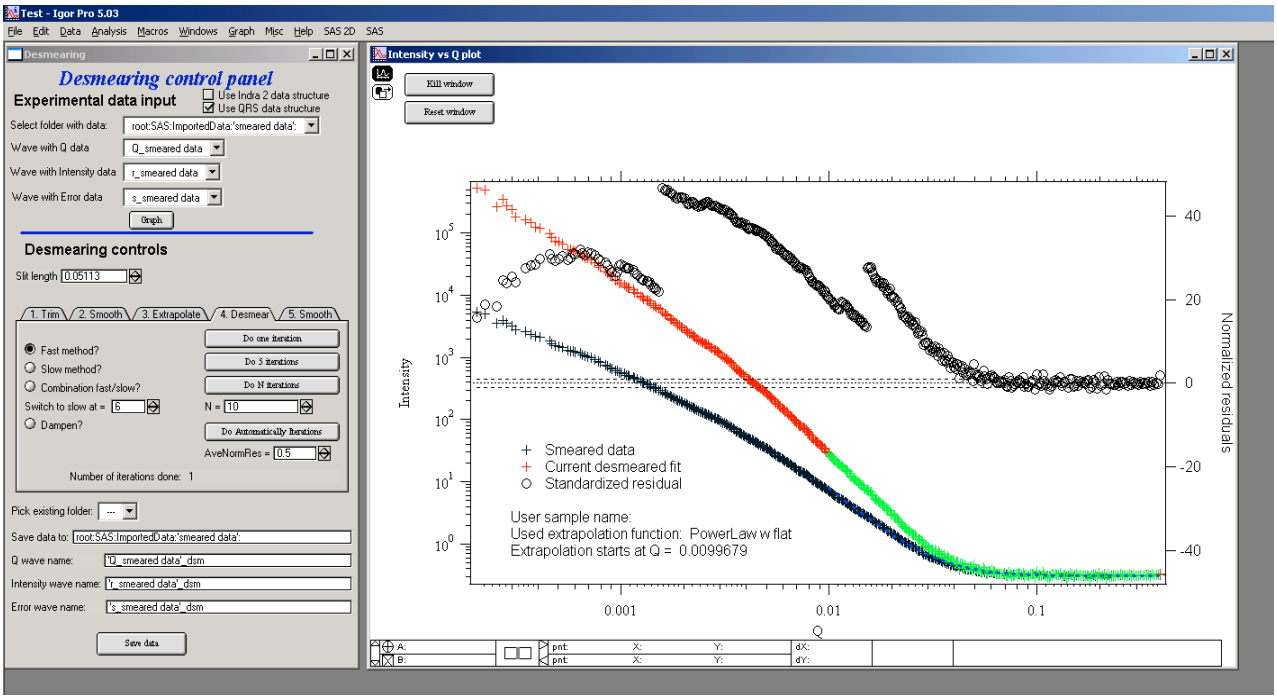
Note the colors: red are the original data, green are the original data used for evaluation of extrapolation parameters and the dotted blue line is the extrapolated data.

4. Desmearing

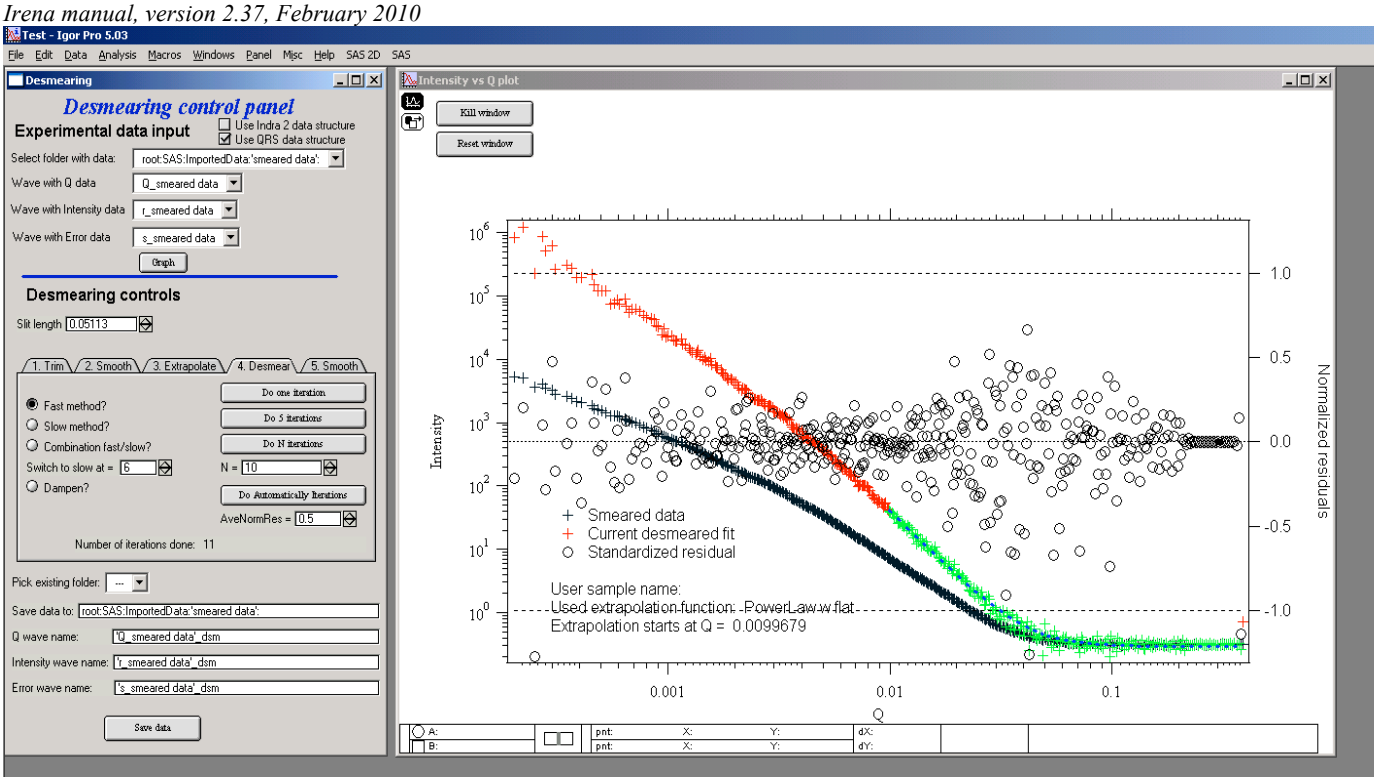


The desmearing can be done in steps – one at a time, 5 at a time, selected number of iterations at once (when you already know how many iterations are going to be needed), automatically (iterates until average normalized residual < preset value) or any combination. Also, there are two modes of conversion for Lake method: aka

"slow" and "fast". The fast method is overall the best method to use, the "slow" method iterates much slower and can result in negative number for intensity.. Combination methods – "Combination", and "Dampen" attempt to use "fast" method (as main) and reduce formation of noise characteristic for this method. In both cases normalized residual for each data point is during each iteration compared. For combination method, if the data point is already estimated to within the user selected precision of input data (normalized residual < User input value) the point is further dersmeared by "slow" method. For dampened method, if the point is estimated to normalized residual < 0.5 it is not desmeared anymore at all…
This should reduce some of the noise created at high-q data during larger number of iterations while keeping the fast convergence of the "fast" method.
Let's select the "Fast nethod" here, for simplicity.
Do one iteration:

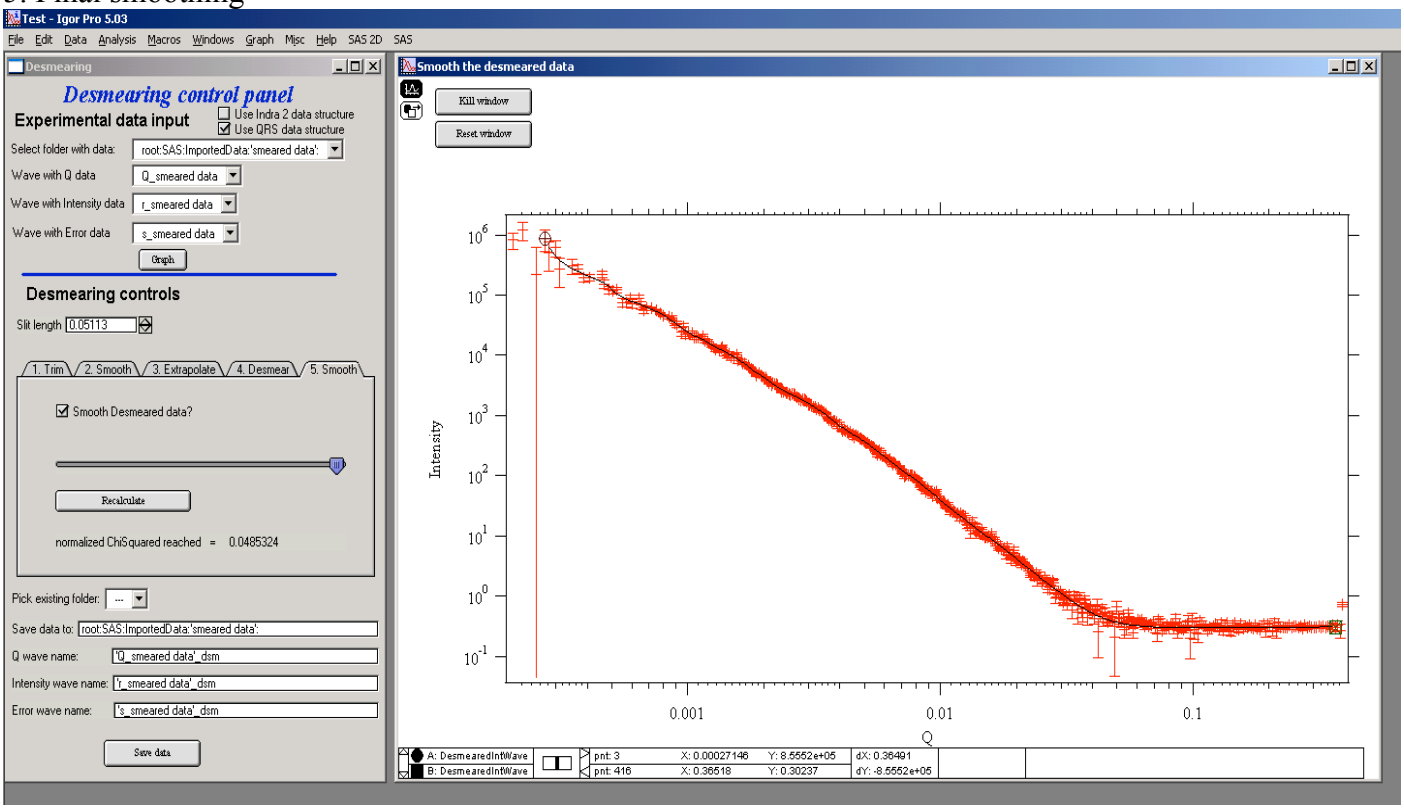


Explanation: Red/green data are current desmeared data (see above about extrapolation). Crosses are original data and circles are normalized residuals.
Desmearing should continue until the plot of the residuals becomes featureless with scatter distributed randomly about z=0 (where z is the standardized residual). Convergence is achieved when the residuals do not readjust to a significant extent between consecutive desmearing iterations. Acceptable convergence is always at the judgement of the person doing the desmearing.



For many data sets, 10-20 iterations are sufficient. Other data sets (those with more structure in the scattering curve) may require as many as 50 iterations or more to satisfy the convergence criteria of the user. For this example data set, this is about where one may end – 10 iterations and most of the circles are within +/- 1. There are some points at low Q which may need more iterations, due to the use of the combination method. (The fast method would have resolved this with fewer iterations.)

5. Final smoothing



Here one can smooth data… This is probably a better place to smooth, if necessary at all.

6. Save data
Use the bottom part of the GUI panel to save data in folder of your choice. The folder, if it does not exist will be created.

7. Unified Fit

7.1 Introduction

Unified fit uses code developed by Greg Beuacage to fit SAS data with levels composed of Guinier part and power law tail. The code handles various data for which development of exact scatterer model is difficult or impossible. It can handle some interparticle interference, fractal scatterers etc. For details, please see papers written by Greg Beuacage or, hopefully, in the future included description.

This is introduction written by Greg for this code:

This fit uses the function described in
http://www.eng.uc.edu/~gbeaucag/PDFPapers/Beaucage2.pdf
http://www.eng.uc.edu/~gbeaucag/PDFPapers/Beaucage1.pdf
http://www.eng.uc.edu/~gbeaucag/PDFPapers/ma970373t.pdf

The basic function is composed of a series of structural levels, each with the possibility to be
a) associated with the previous smaller size level (Rcutoff2 = Rg1 in I2highq=B2q^(-p2)exp(-q^2Rg1^2/3) for the power-law region of 2)
b) to follow mass fractal restrictions (calculate B for the mass fractal power law I = B q^(-p)
*c) to display spherical Correlations (Interference) as described by I(q) = I(q)/(1+p f(q etai)) where p is a packing factor 8*vH/vO for vH = hard sphere volume and vO is occupied volume and f(q eta) is the sphere amplitude function for spherical Correlations (Interference)*

The intensity from each level is summed and the intensity from one level, i, is given by:

Ii(q) = Gi exp(-q^2Rgi^2/3) + exp(-q^2Rg(i-1)^2/3)Bi {[erf(q Rgi/sqrt(6))]^3/q}^Pi

This equation includes a) above if Rg(i-1) is the previous smaller Rg e.g. the primary particles from a mass fractal level. If there is no such dependence Rg(i-1) is set to 0 or it could be set to an independent size under unsual circumstances

This equation can include b) if Bi is calculated using Bi = (G df/Rg^df) GammaFun(df/2) and the erf argument includes kqRgi/sqrt(6) where k is 1.06. The latter can be included or ignored for high dimension mass fractals but becomes more important for dimensions less than 2.

The equation can include c) by multiplying the entire level Ii(q) by a function that follows the Born-Green approximation for Correlations (multiple particle Correlations) and this works well for weak Correlations of any type but becomes more restricted to spherical Correlations as the Correlations become stronger. The measure of the strength of the Correlations is the packing factor p = 8 vH/vO as described above and for spherical particles this value can be 0 (no Corelations) to about 5.92 (calculated for FCC or HCP packing).
*The packing factor for FCC is p=8*Vh/Vo = 8*(pi*sqrt(2)/6)=5.92 - this is basically the total volume of latticepoints in FCC devided by the volume of the lattice (provided to me byone nice Irena user)*.
If asymmetric particles (rods or sheets) are packing the number can be much higher and the spherical function becomes less appropriate although it can be used in a pinch for weak Correlations. The interpretation of p and eta become complicated in these cases. As a general rule etai has to be larger than Rgi as common sense would dictate. The correlation function follows closely the development of Fournet in Guinier and Fournet and in Fournet's PhD dissertation where it is better described but is in French...

So the Unified needs to accommodate multiple levels each of which can potentially have 8 parameters (including spherical Correlations):
Rgi,
Gi,
Pi,
Bi,

etai,
packi,
RgCOi,k where RCOi is usually Rg(i-1), as shown above, for hierarchical structures (k is 1.06 for mass fractals and 1 for others)
Each level must also have the answer to at least three questions:
Are there Correlations: qCori.
Is this a Mass Fractal: qMFi.
Does this level terminate at high-q in the next lower level Rg: qPL (PowerLimit) That is, is this a hierarchical structure build from the previous smaller level. A third option is to let the power law limit float as a free parameter although this is rarely appropriate.

As Gregg wrote me:
"We have several options for coding the unified function.

a) Write a dedicated code for a specific morphological model where all of the parameters are defined in terms of the model. We have done this for correlated lamellae, rods, mass-fractals, spheres, correlated spheres, RPA based polymer blends of arbitrary fractal dimension, polymer gels among others.

b) Write a generic unified code that allows a high degree of flexibility but which is naturally complex. For cases where you deal with a fairly complex and limited structural model option a) is most appropriate and is easiest to understand. We can't however write such code for each and every case. Several of our publications indicate how to go about calculating the unified parameters, for instance for a sheet structure 8 parameters in the unified equation (for 2 levels) reduce to 3 free parameters, the contrast, thickness and diameter of the sheets. Similarly rods can be described by 3 parameters the length, diameter and contrast. Correlations in both systems add 2 other parameters although the spherical correlation function can not be rigorously used except at extremely weak levels of correlation."

This code deals with approach b) where only spherical correlations are dealt with but including an optional mass fractal limitation (strictly limited to linear chains but useful for branched structures in application).

Helpful hint on use of Invariant:
The invariant INV is:

$$INV = 2*pi^2*\Phi(1-\Phi)*\Delta\rho^2$$

To use correctly, one needs to convert the Unified provided invariant to cm^{-4} by multiplying by 10^{24}
(from $cm^{-1}A^{-3}$ to cm^{-4})….

Helpful hint on use of Porods law:
The specific surface area S_v is:

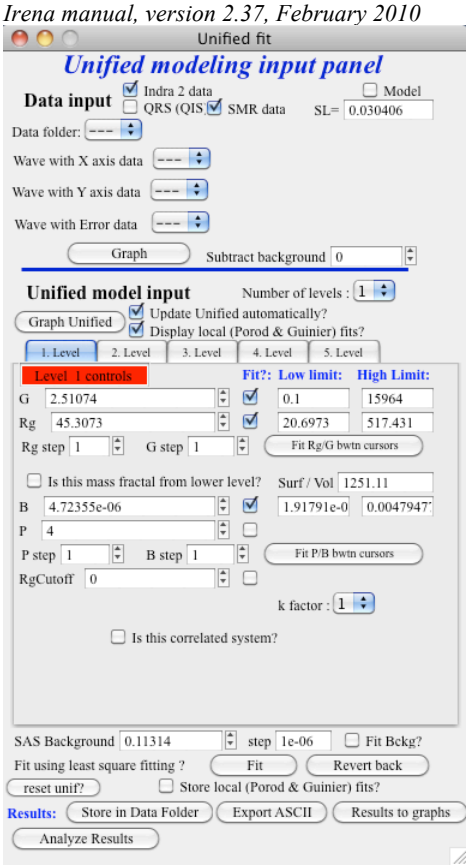
$$S_v = B / 2*pi*\Delta\rho^2$$

Where B is Porods constant – it is the value you get from Unified fit when P=4. To use correctly, one needs to convert the Unified provided B to cm^{-5} by multiplying by 10^{32}
(from $cm^{-1}A^{-4}$ to cm^{-5})….

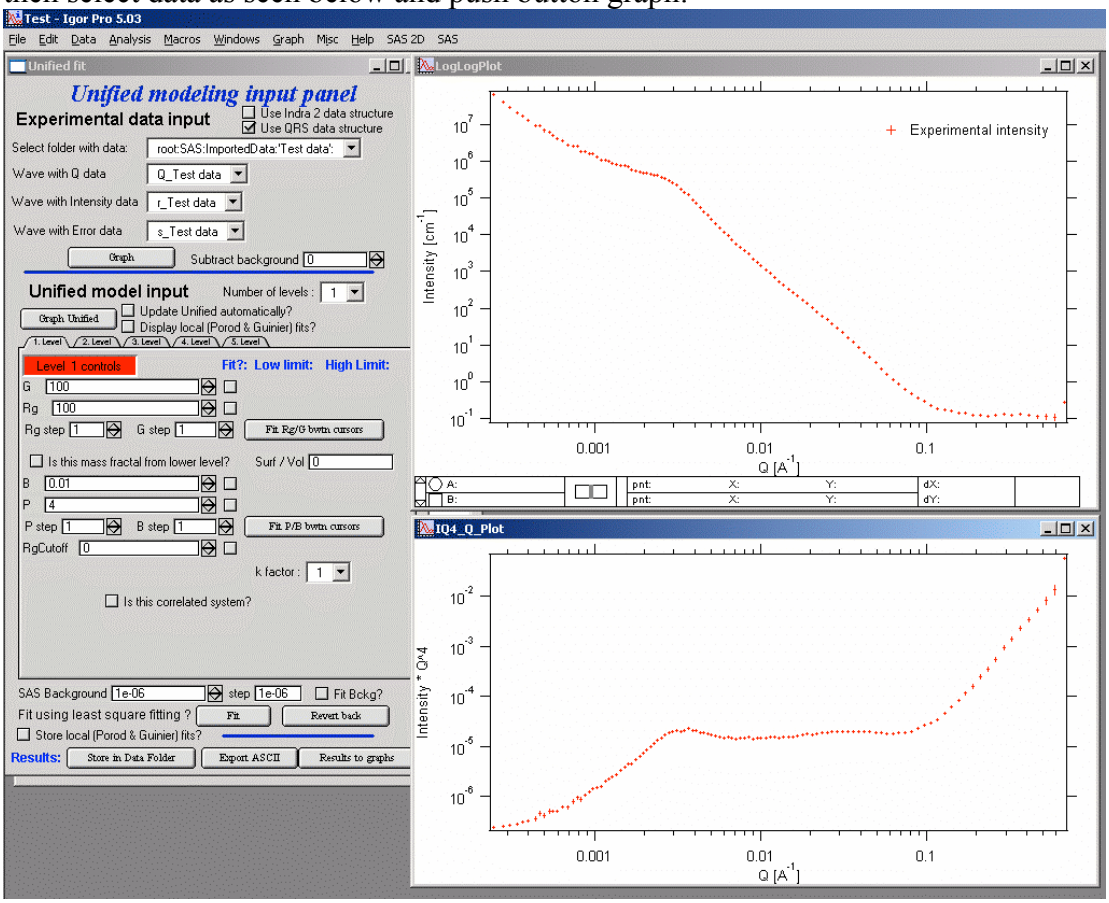
NOTE: from version 1.37 Unified has "Analyze results" tool, which can analyze also Invariant, Porods law and some other specific cases.

7.2 Running Unified fit macro

Select "SAS" – "Unified fit model". Following is the screen you should see after initialization:



In the top part again deselect “Use Indra 2 data structure”, or “Use QRS data structure” or none checkbox and then select data as seen below and push button graph:



The two graphs which appear show selected data in two different views – top graph is log-log Intensity vs Q vector, bottom part is Intensity * Q^4 vs Q vector, this represents S(Q) as known from scattering theories.

ALL USER INPUT IS IN THE TOP WINDOW (log-log plot)

Controls above the tabs:

Graph Unified button — recalculates model and puts it into the graphs

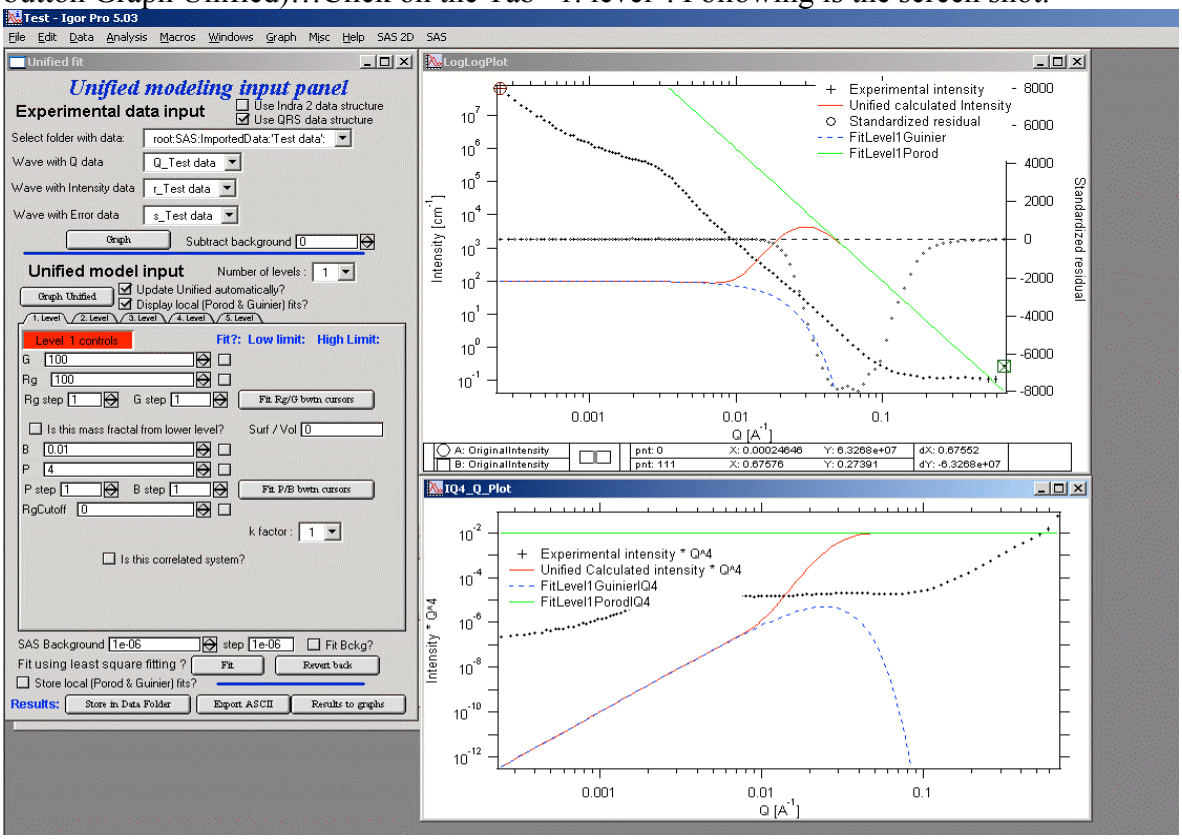
Update automatically checkbox — if any change to model parameters is made, automatically recalculates (i.e., automatically pushes the button Graph Unified). Uncheck on slow computers.

Display local fits — displays local Guinier and Porod fits in the graphs for active tab level

Number of levels — select number of levels displayed. Note, that at no time parameters for not displayed levels are reset, so user can switch between number of displayed (=used) levels freely…

Now, when we have the data we will start building the structures from large Qs…

Select number of level 1, check Update Unified automatically (if used computer is fast enough, else push the button Graph Unified)…Click on the Tab "1. level". Following is the screen shot:



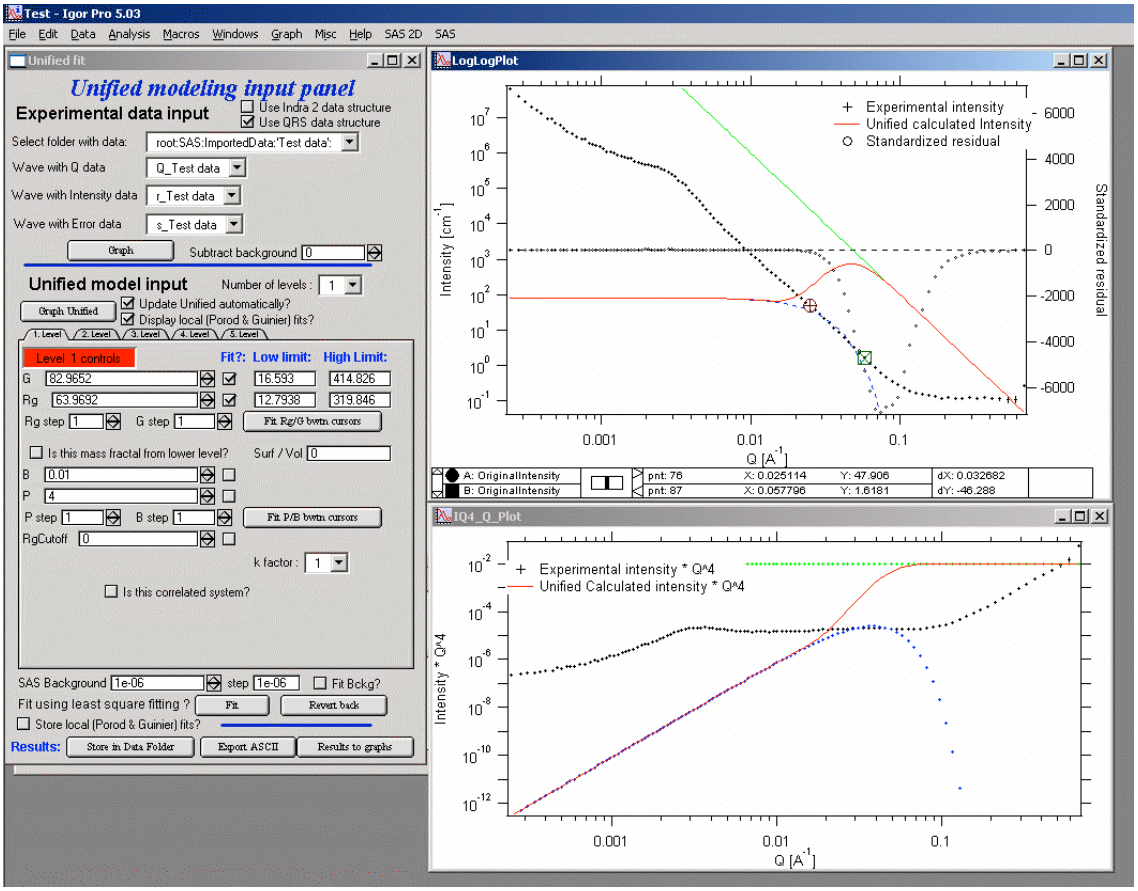
Description of Tab area in the panel:

This area changes according to parameter selection etc. I have tried to put in this as much "smarts" as possible to help user make right choices, so parameters, which are not applicable at any given time should be invisible… Anyway, each parameter has most controls grouped on one line – that is current parameter value, checkbox "Fit" and (if Fit is selected) limits – low and high fitting limits. Further underneath the parameters is parameter step variable. This changes the step with which given parameter changes when arrows at the end of it's field are pushed (up or down). This allows user to "walk" the parameters into their starting condition as best as possible. There are also buttons for local fits. If checkbox "Is it correlated system" is selected, new parameters appear. Also screen changes, if the checkbox "Is it mass fractal from lower level" is selected….

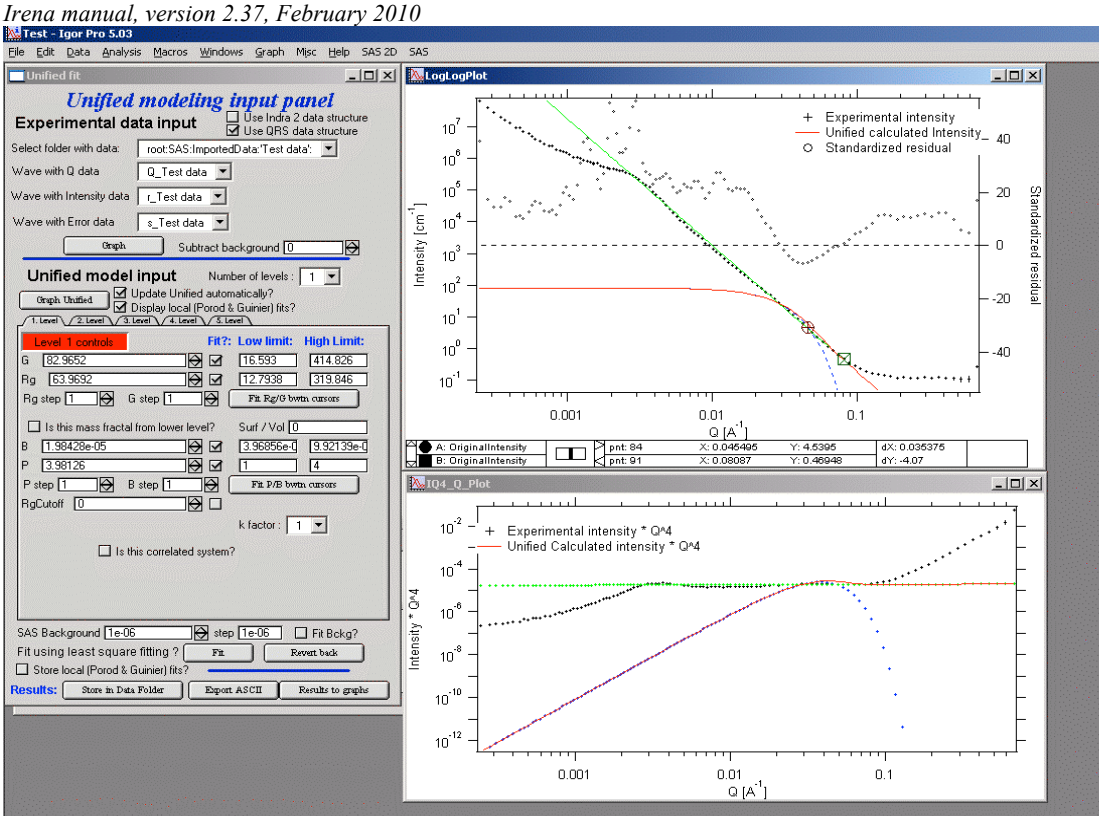
The new red line in the graphs represents the results f the model for default values for level 1. We need to now make this level to fit the smaller particles Guinier and Power at Q ~ 0.025 A^{-1}. This shoulder is better visible in the bottom graph. Also we need to include appropriate background…

Change background (under the tabs field) into the 0.1

Select point 77 –86 on the top graph using cursors and check "Fit" checkboxes next to G and Rg. These parameters will be fitted in between the cursors. Push button "Fit Rg/G bwtn cursors". Following should be result:



The blue line in the graphs now is the Guinier fit. Next select points 84 to 92 with cursors, check boxes "Fit" next to B and P and push button "Fit P/B btwn cursors". Now we get the powerlaw fit on this area:



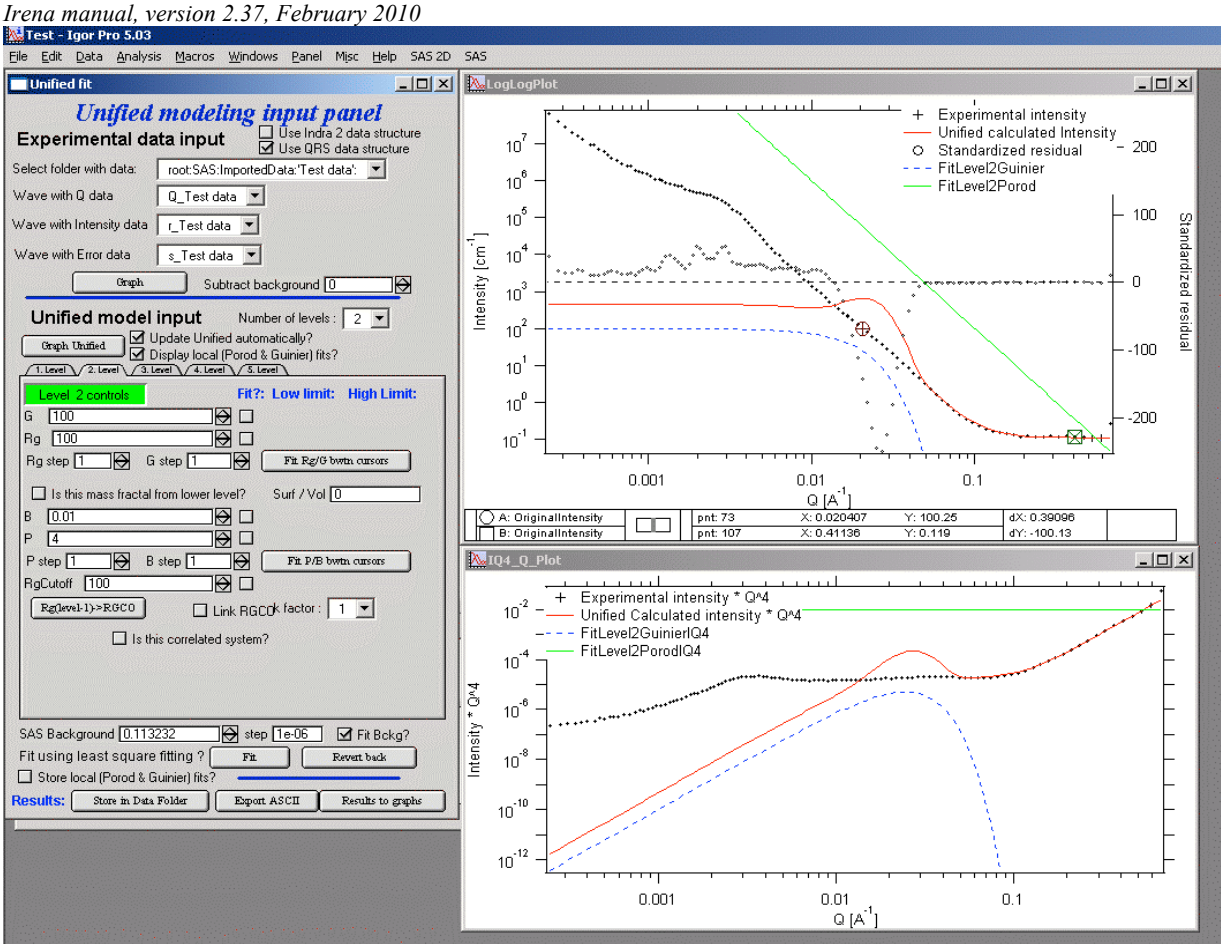
The green line is local fit for the power law dependence. Notice, that this fit is now reasonably good, including our background estimate.

You can also guess flat background, or check the checkbox “Fit Bckg?” and let least square fitting (next below) to estimate best guess for flat background.

Now we can fit the whole Level 1 dependence. Select large range of data (points 77 – 109) and check “Fit” next to background parameter. Then push the fit button. This runs least square fitting on the data.

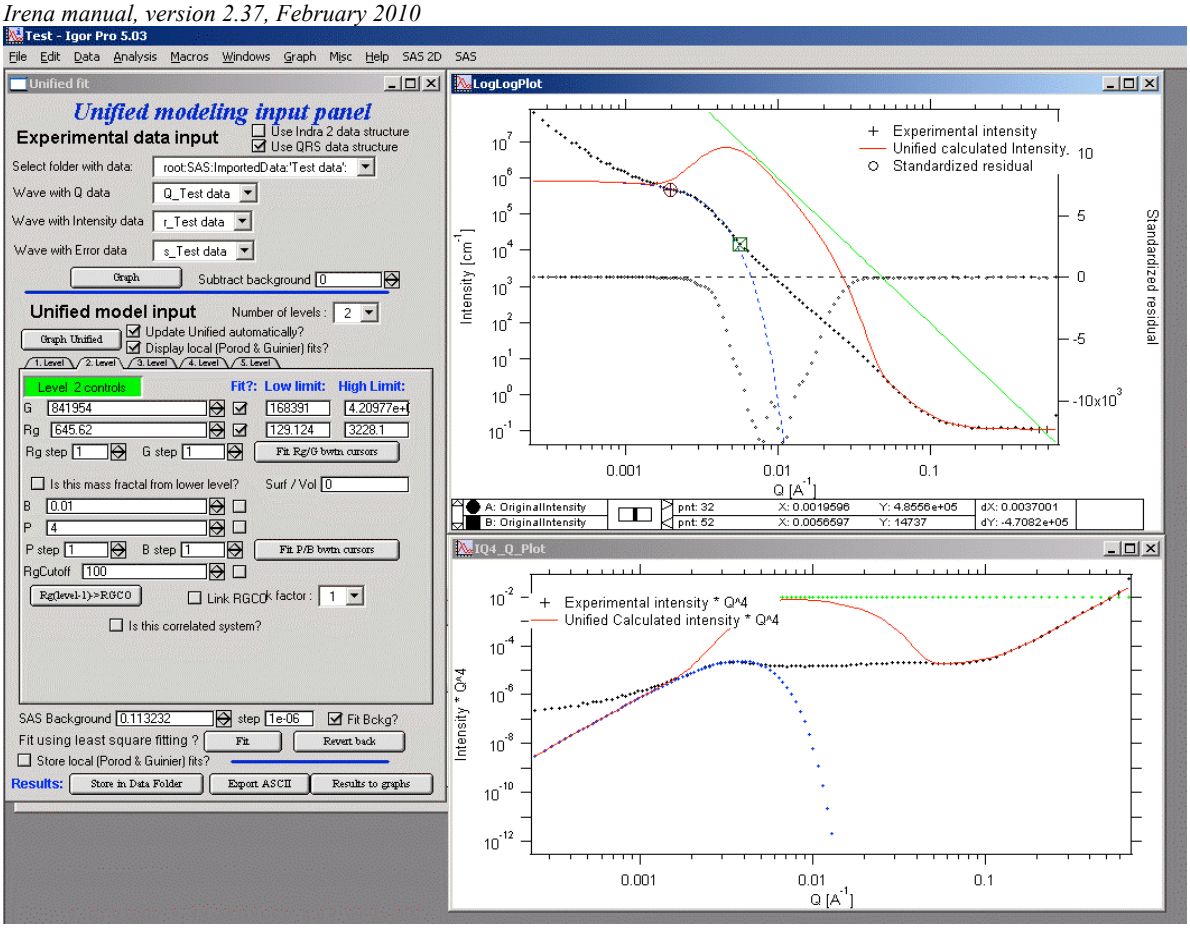
Comment:
If the least square fit fails, it should automatically return all values where they were before fitting. If the fit walks away to solution, which is not right, user can push button “Revert Fit” and parameters are returned back where they were before fitting.
Only parameters selected to fit (checkbox “Fit” next to them) are fitted – **BUT FROM ALL ACTIVE TABS**. Therefore if using more than one level make sure you have selected only appropriate parameters from all levels you want to vary in this fit. ***These fits can be highly unstable, if starting conditions are not right***….

Ok, level 1 looks fine and the background also. Uncheck all fit boxes in the panel and then select Number of levels as 2. Click on tab with “2. level”.

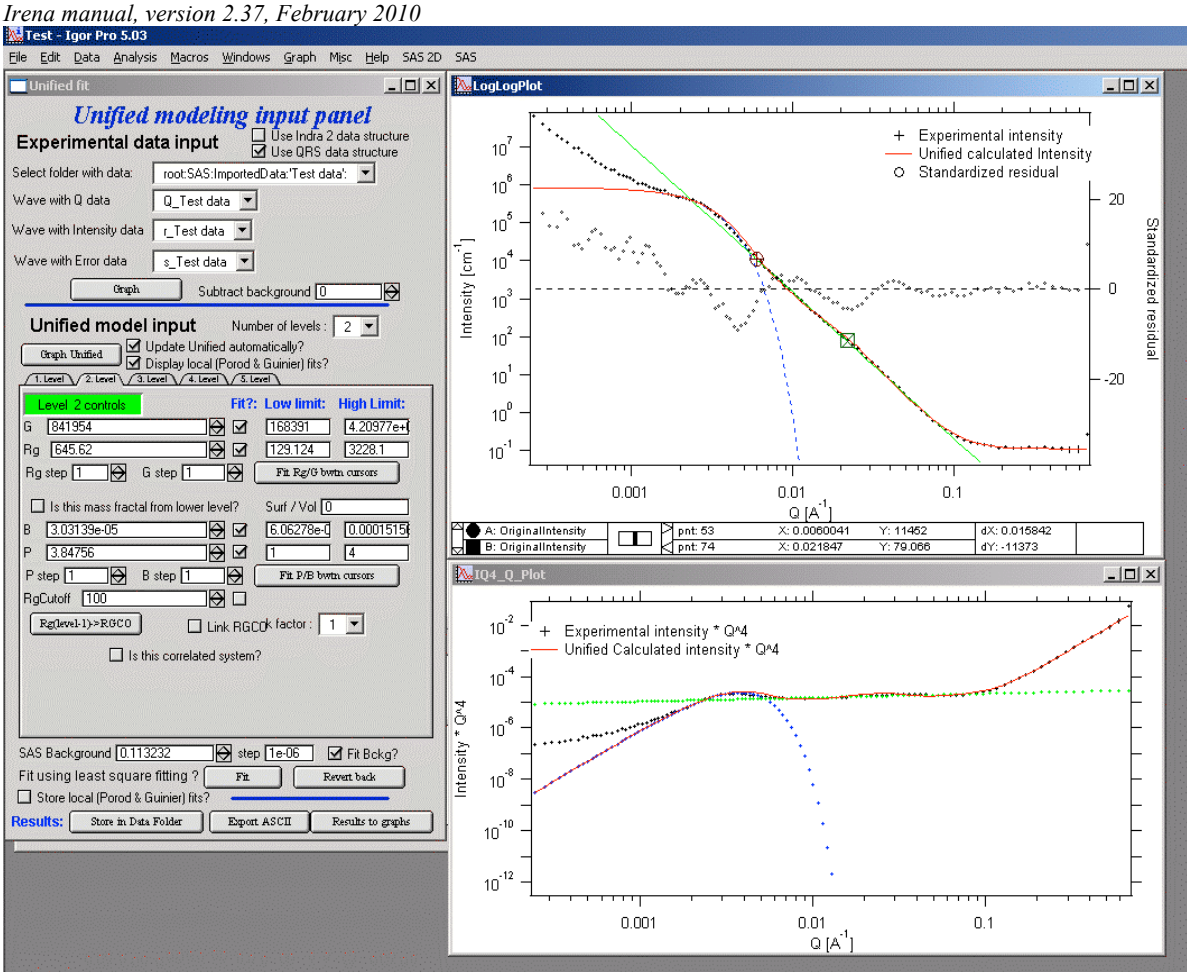


The green and blue lines in at this time represent default values for level 2. If you click now on tab of 1. level, the blue and green lines will be set for local fits for level 1, if on tab with level 2, they represent local fits for level2., etc…

Let's do local fits. Set cursors between points 31 and 48. The default values for the parameters are way to far from the expected values and the fit may not converge from them, so change G significantly higher (just add few 0 to make it may be 100 000 or more) and increase the Rg also to some higher number (500 or so). Note, that the blue fit line moves closer to the measured data. Next push "Fit Rg/G bwtn cursors" button. Fit should converge to following solution:



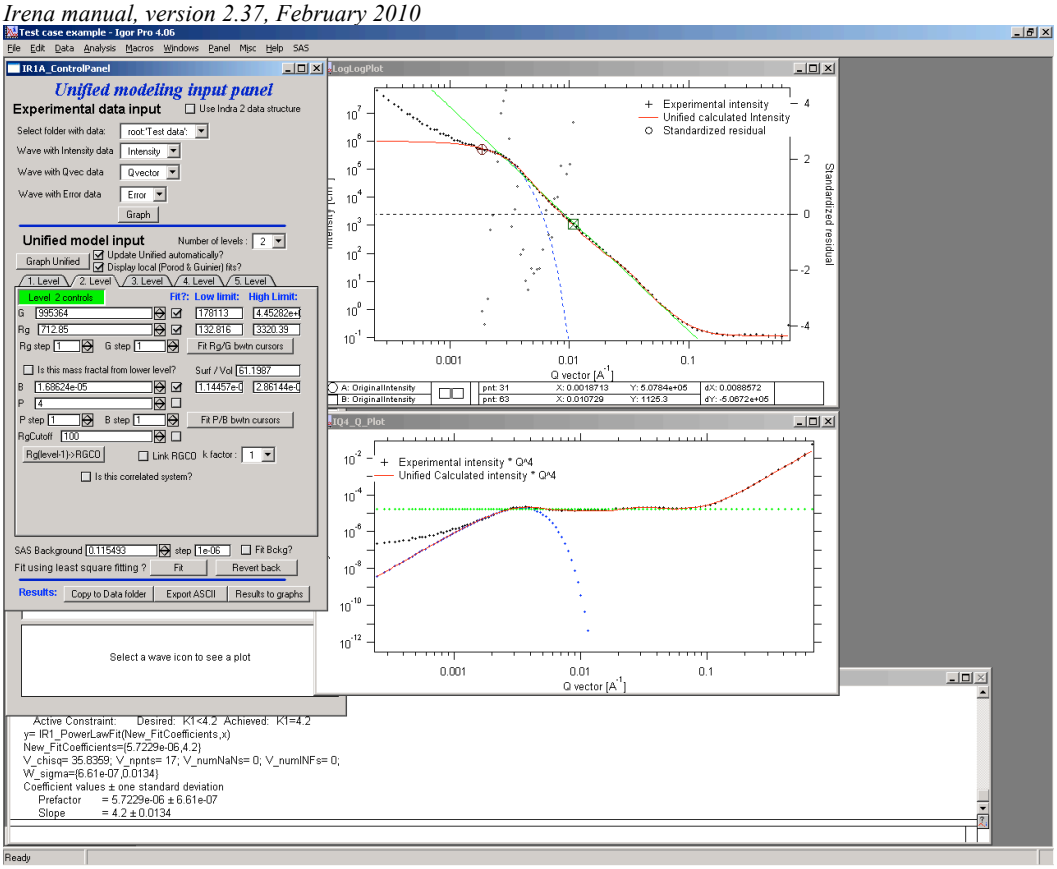
Now we need to do something similar for power law dependence. Select points 47 to 63, reduce B to about 0.00001, select “fit” checkboxes there and do Local fit. Following should be the result:



Now we can select area with this level only and optimize the parameters of Both Guinier part and Power law part together:

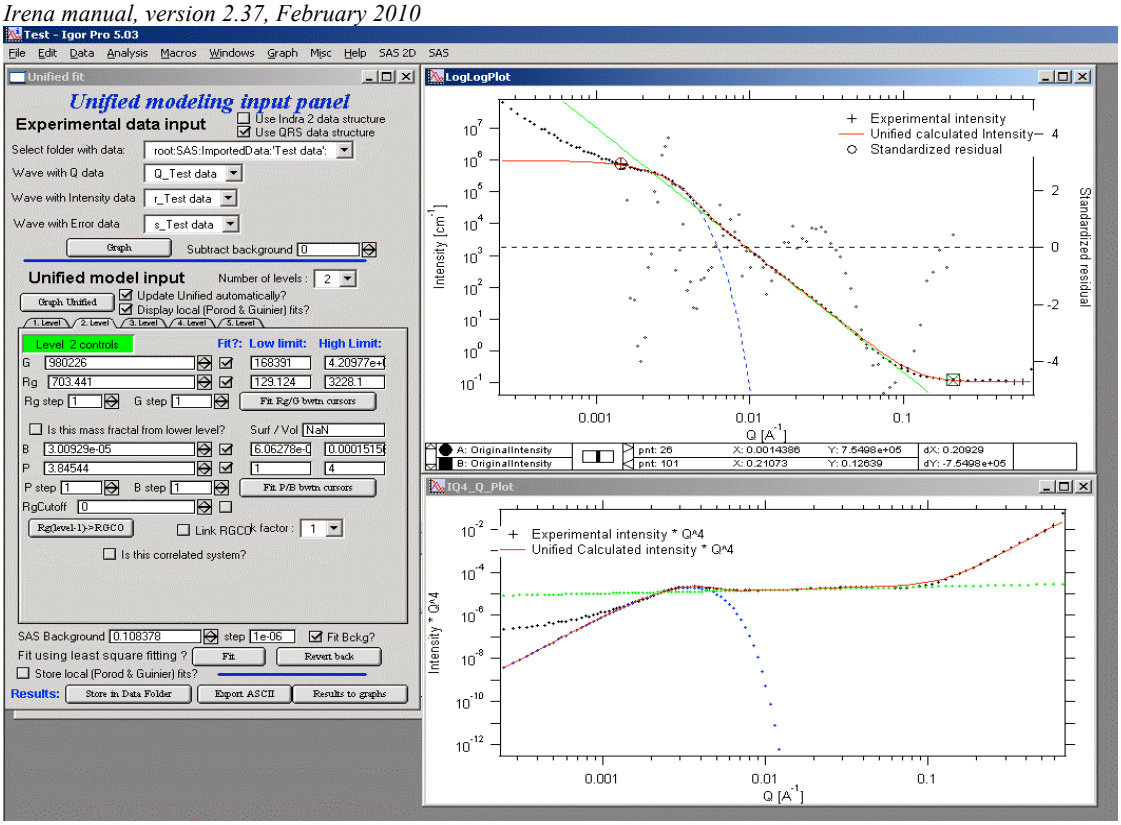
Select point 31 to 63, make sure that other levels (level1) parameters have deselected checkboxes Fit, and push fit button. If you get error message that limits are incorrectly selected, check, that starting conditions for that parameter are between the limits. This is very important…

In this case the reason for error message is the fact, that power law slope starting value is 4.2, which is not physical. Change that to 4 and fix it by deselecting the Fit checkbox. The do the fit:

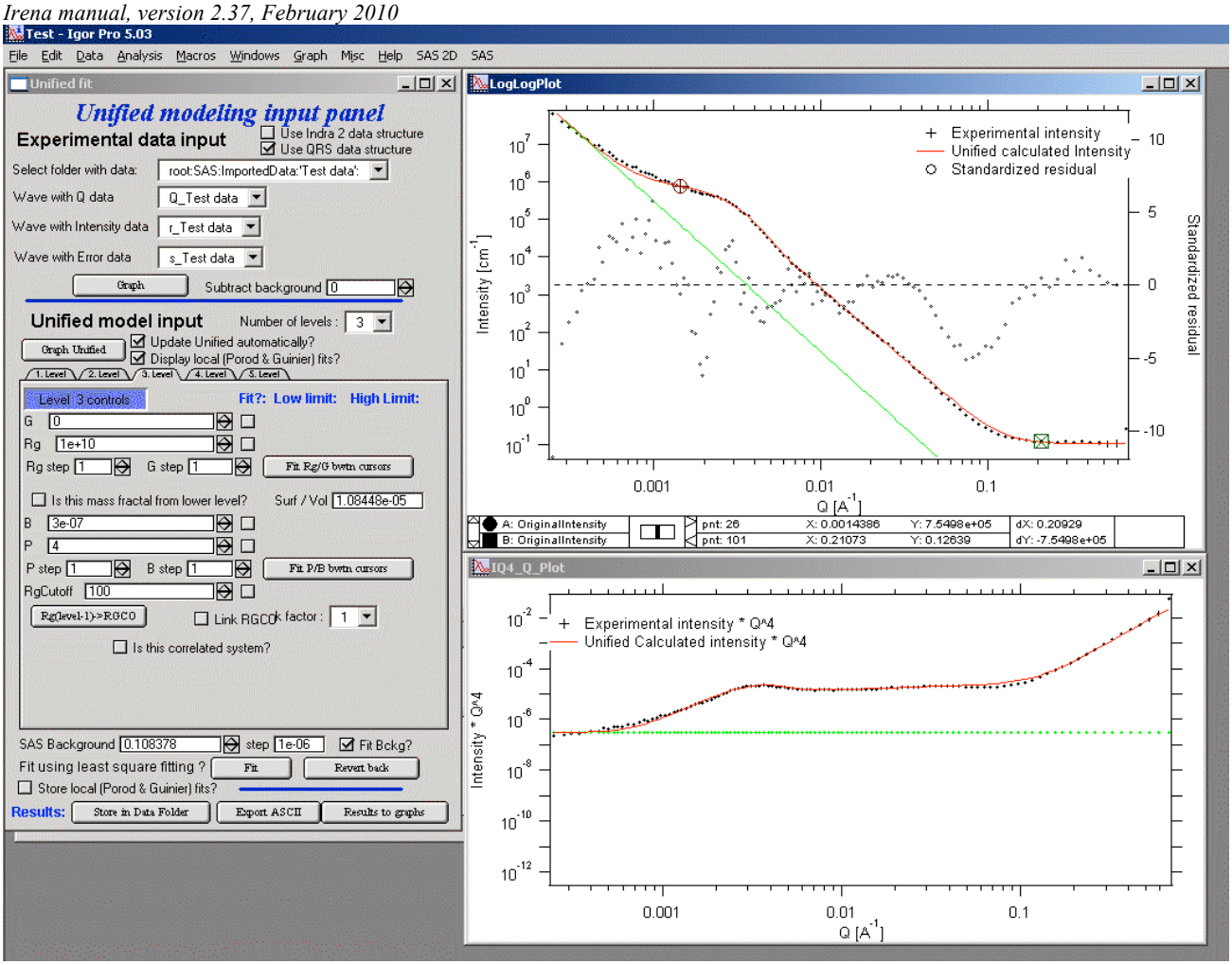


And now is the time to broaden the data range and fit both levels together. It is not necessary to fit background, but should not hurt… Select points 31 to 103, check all parameters for level 1 and level 2 to fit and then fit.

Note, that in this case (this was mixture of two powders) the right setting for Level 2 (large particles) RgCutoff (RGCO) parameter is 0, since the scattering from these large particles extends to even largest Qs. Therefore, to get good fit one needs to set RGCO to 0 and rerun the fit. Some modification of starting conditions may be necessary (I had to set B for level 2 to lower number to get stable solution). But one can get really good solution:

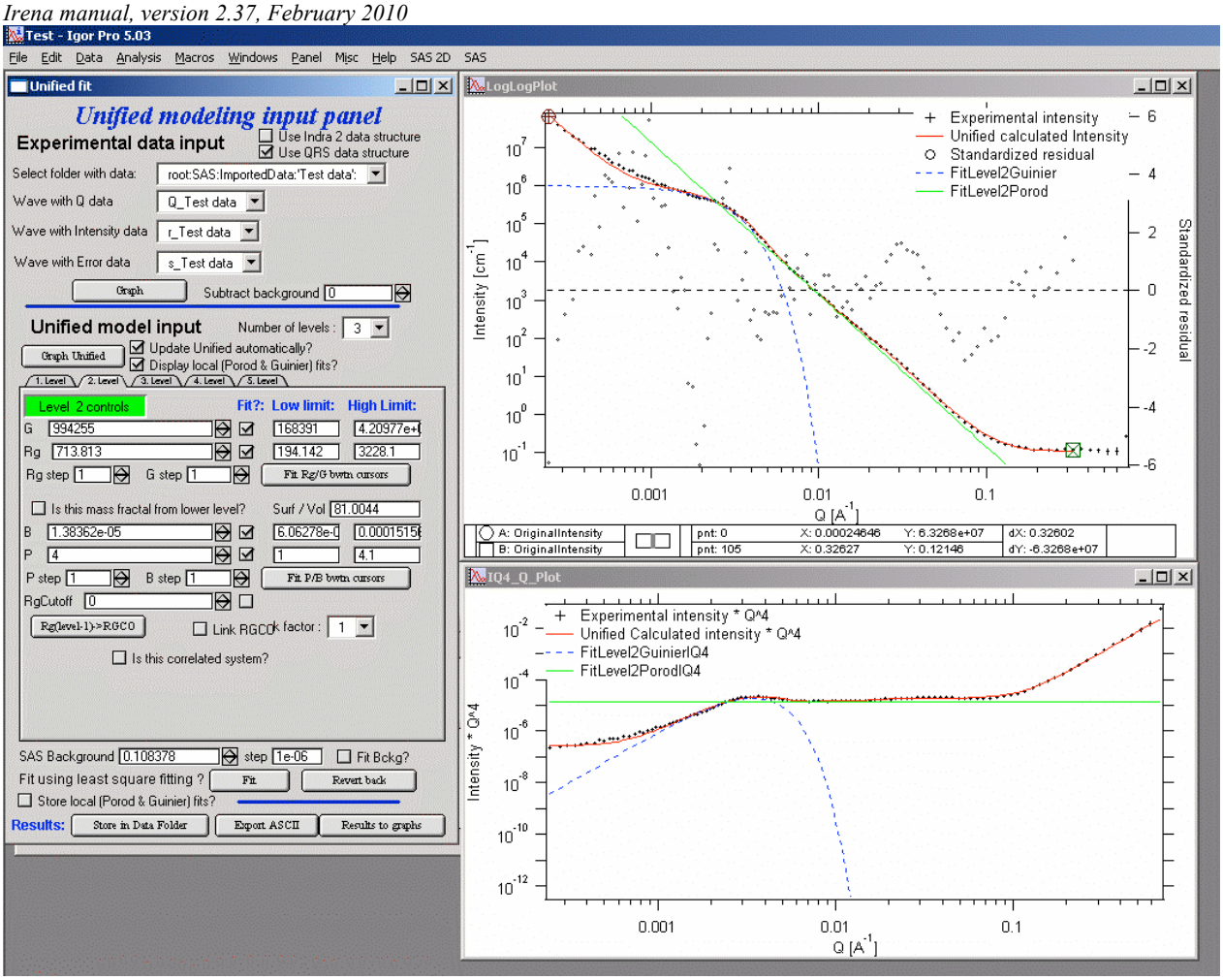


Last part, which may needed to be solved is the power law tail at smallest Qs from something large. To do so, we can use trick of this method – create population, which has really large Rg, but G is 0, This removes effectively Guinier area from the calculations. Then we are left with power law only. To do so, create level 3 and select the tab with level 3. Set G to 0 and Rg is automatically set to $1 * 10^{10}$. Now we need to fix only the powerlaw part. This is best done by modifying the B and P manually… A good guess clearly is about $3*10^{-7}$ for B and P roughly 4.



Now select wide range of Q's – points 103, select parameters to fit (possibly all, but that will take longer) and make sure the limits (especially for level 3!!!) are set correctly. Note, that Level 3 Guinier parameters should not be fitted! Then push button fit.

And we receive nice solution:



notice the standardized residuals in the top graphs are reasonable for all Q's suggesting that we have right number of levels. It may be possible to improve the fit by including some correlations – the powder could have been compacted quite a bit, but I leave that to user to figure out more…

7.3 Correlations

If interparticle interference is not negligible, then for reasonably weak interferences the code has built in simple model for modeling those. This is simple model which is realistically valid only for gasses and is only approximation. For details see for example …..

User should be aware of the crudeness of these calculations.

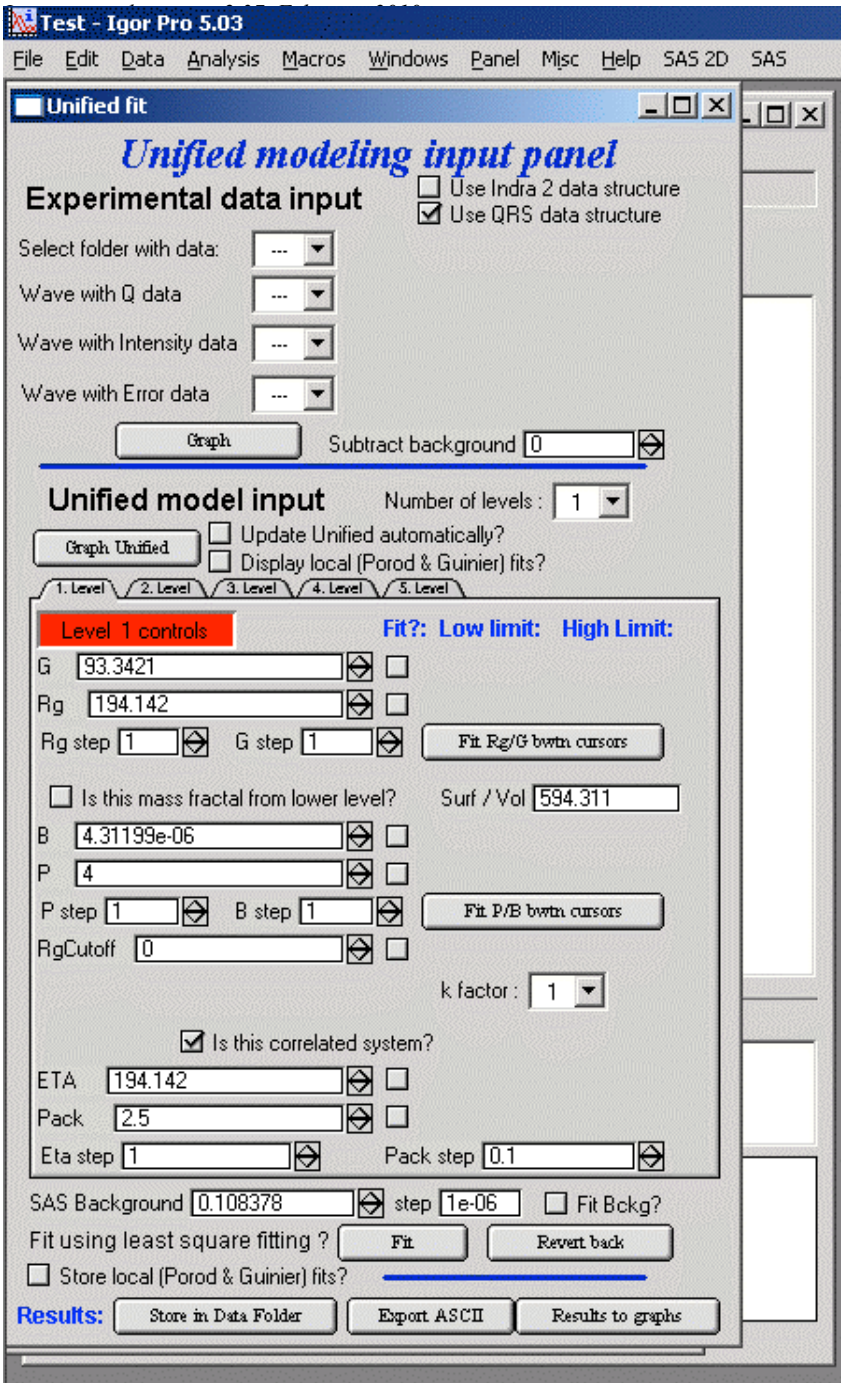
The code used for calculations involves correcting intensity from a level using this formula:

$$\text{Intensity}_{\text{with interfernce}}(Q, R) = \text{Intensity}_{\text{without interference}} * (1 + \textit{pack} * \text{SphereAmplitude}(Q, \textit{Eta}))$$

Where the *pack* and *Eta* are the two parameters of this model.

Note, that this is supposed to be valid for spheres. I am working on adding other methods for other arrangements of particles.

Remember: this method accounts in very crude way ONLY for interaction for particles in the particular population. If there are interactions among particles from different populations – which is very likely – these calculations have NO WAY to account for it.



When checkbox is selected for correlations, new windows appear – ETA (distance between the layers) and Pack (fill of the first layer). Smaller the Pack, less interference. The ETA should not ever bee smaller than size of particles, and actually should be larger…

7.4 RgCo again – main warning

My experience has shown, that one of the least understood parameters of the whole Unified fit seems to be RgCo parameter. Here is more details on this parameter:

If you look in the formulas and what this parameter actually does, you will see, that it terminates very steeply scattering form given population by the time the one reaches Q ~ RgCo. Therefore the level becomes unimportant at q higher than equivalent of RgCo.

There are two cases when one needs this parameter and both relate to case when higher level and next lower level represent scattering from the same volume of materials.

1. Scattering from particles having two main dimensions – such as rods, disks etc. In this case the form factor (see the pdf list of form factors) exhibits two Guinier regions connected by relatively shallow power law slope. After the higher-q Guinier are the terminal slope is Q^{-4}. In order to be able to describe this type of behavior the higher level power law scattering MUST be terminated by the time we reach the lower level Rg.
2. Scattering from fractals which exhibit more than one characteristic dimension in the measured Q range. The argument here is VERY similar. Imagine fractal measured over such q range, that one can see the fractal behavior (higher level) but also time when you can see the primary particles. This very much resembles the case 1, except it is less clear.

My general simplistic rule is, that if the two levels represent scattering **FROM ONE POPULATION (VOLUME) OF PARTICLES** then the RgCO must be set, if these are different populations (having their own volumes) then the scattering is additive and RgCO should be set ot 0 for both of the levels.

7.5 Output from Unified

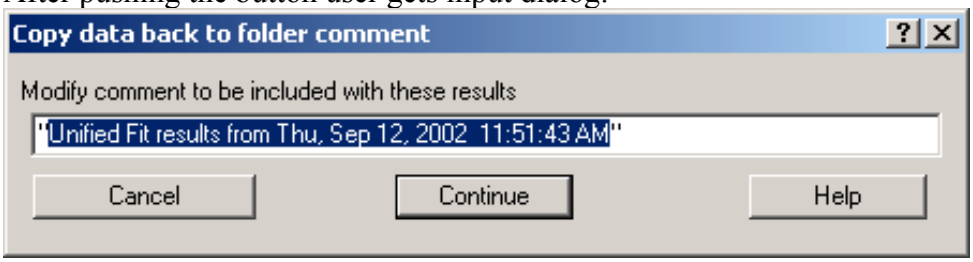
Result scan be either copied back to folder where the data came from, exported as ASCII, or little macro will include for each level text box in both of the graphs. User than can modify fonts/size etc and print. I need to make this later more user friendly to give more flexibility…

Further the data can be exported into Excel table using "Export to XLS file panel". This macro was written by Greg Beaucage and I need to learn myself little bit better what it does. But it allows output results into spreadsheet for publication.
Use the buttons at the bottom of the panel.

Copy to Data folder

If checkbox "Store Local (Porod & Guinier) fits?" is selected, then saved are not only final fits but also all local fits too.. This creates large number of waves, but provides separated outputs for various levels – allowing to use these data for further separate analysis…

After pushing the button user gets input dialog:



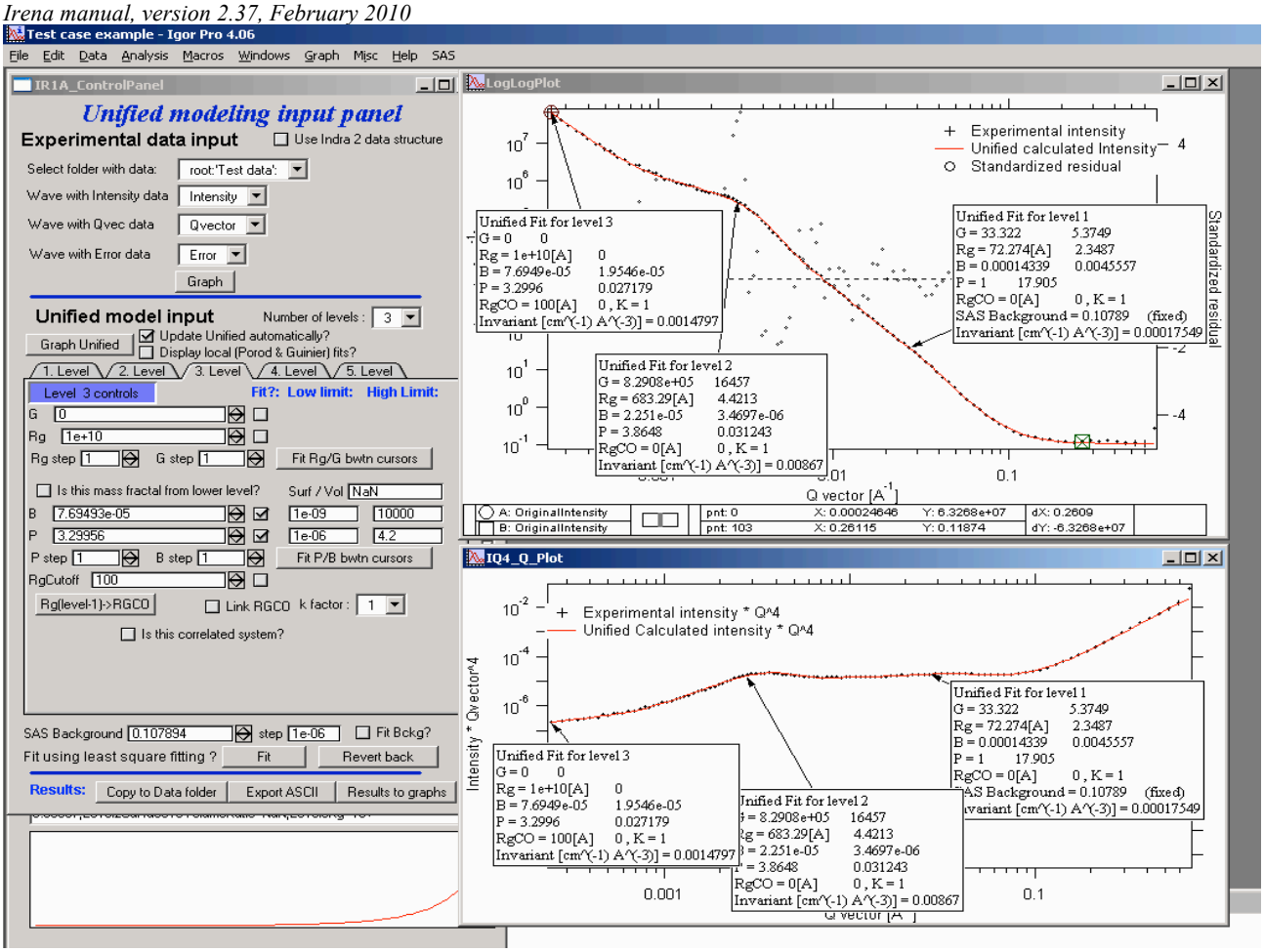
In which any useful comment can be inserted (modify default). Note the quotes. They have to be there…
Then program saves following waves in the folder with original data:
UnifiedFitIntensity_0
UnifiedFitQvector_0
The _0 is generation number. User can save large number of solutions, with increasing _XX where XX is number. When Unified is run on data in folder, where Unified solution exists, user can recover any present solution – all parameters are put back in the panel, this allows user to quickly return back to previously saved solution, whit out need for recording the results.
All Unified fit parameters are saved in the wave notes of the above listed waves. This list is quite extensive and hopefully the names are descriptive enough. User can interrogate them either in data browser or using Igor built in tools (read functions "note", and "StringByKey" resp "NumberByKey" manual)…
This is the example of the list in this case:
*IgorExperimentName=Test case example;DataFolderinIgor=root:'Test data':;DistributionTypeModelled=Number distribution;NumberOfModelledDistributions=2;SASBackground=0.12257;Dist1ShapeModel=sphere;Dist1ScatShapeParam1=1;Dist1ScatShapeParam2=1;Dist1ScatShapeParam3=1;Dist1DistributionType=LogNormal;Dist1Formula=P(x)=(1/((x-loc)*scale*sqrt(2*pi)) * exp(-ln((x-loc)/scale)^2/(2*shape^2));Dist1NegligibleFraction=0.01;Dist1VolFraction=0.0024333;Dist1Location=21.79;Dist1Scale=87.731;Dist1Shape=0.5;Dist1VolFractionError=8.698e-05;Dist1LocationError=16.734;Dist1ScaleError=5.8733;Dist1ShapeError=0;Dist2ShapeModel=sphere;Dist2ScatShapeParam1=1;Dist2ScatShapeParam2=1;Dist2ScatShapeParam3=1;Dist2DistributionType=LogNormal;Dist2Formula=P(x)=(1/((x-loc)*scale*sqrt(2*pi)) * exp(-ln((x-loc)/scale)^2/(2*shape^2));Dist2NegligibleFraction=0.01;Dist2VolFraction=0.047415;Dist2Location=608.88;Dist2Scale=538.71;Dist2Shape=0.5;Dist2VolFractionError=0.00026279;Dist2LocationError=13.656;Dist2ScaleError=7.249;Dist2ShapeError=0;UsersComment=Result from Modeling Thu, Sep 12, 2002 1:20:06 PM;Wname=ModelingQvector_0;Units=A-1;*

Export ASCII

This exports ASCII file with all the fitting parameters from Unified model (whole wave note from the results waves) and Q and Measured intensity, Error estimates and Unified fit intensity.

Results to graph

This includes results in the graphs in text boxes:



Note, that only selected, applicable parameters for each level are included in the text boxes… The text boxes can be formatted (double click on the text box) to suit user output. But note, that the text boxes get redrawn (and therefore reset to default) next time user pushes the "Results to graph" button.

Export to XLS file panel

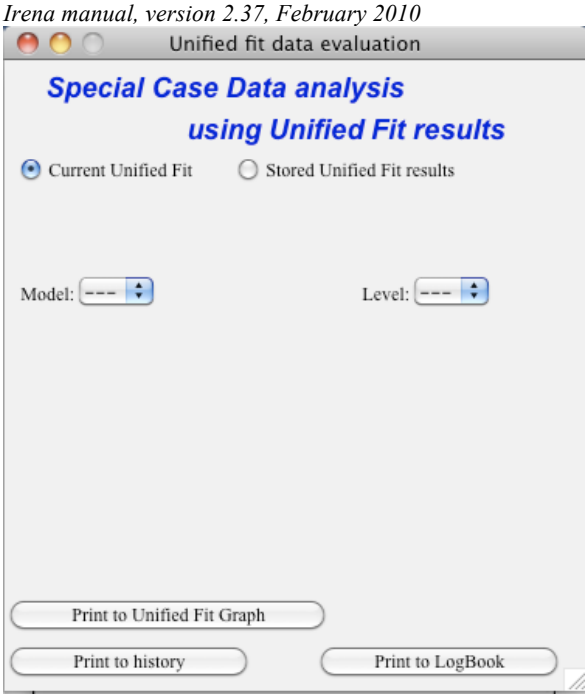
This macro should output data in the table which can be loaded by spread sheet type program (Excel). Macro creates panel with buttons and walks user through steps needed to add data to the notebook in Igor, which then can be saves as text file and imported to other programs.

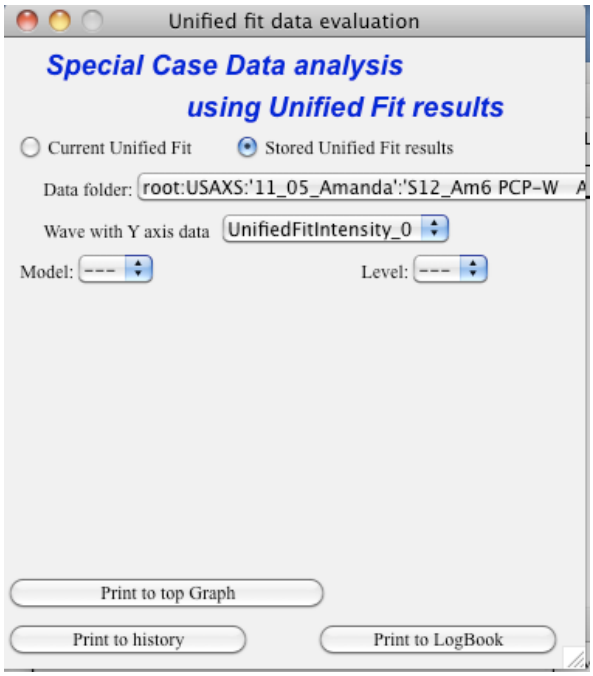
7.6 Analyze results

Some specifc cases can be analyzed further using Unified method. These are :
Invariant
Porod's law
And special cases publisched by Greg beaucage of:
Size distribution and Branched polymers.

All of these can be from version 1.37 analyzed by using "Analyze results" tool. It can be called from the bottom of the Unified main panel.

Usage:
Open the tool from the Unified panel (at the bottom):



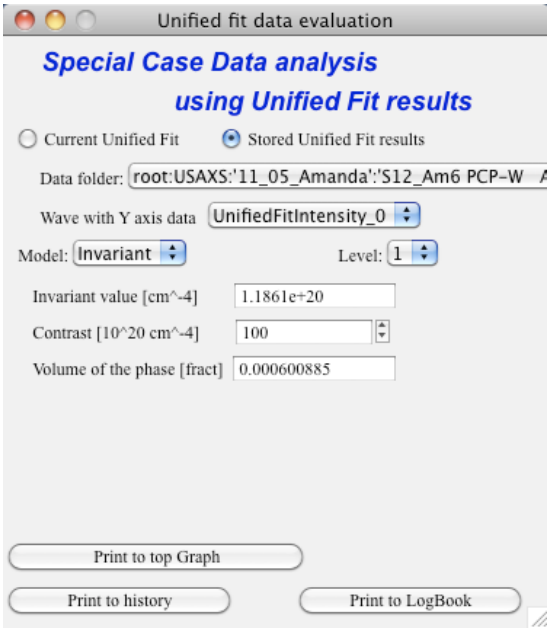


There are two options which data can be analyzed…
Current Unified data in the Unified fit tool. This is selected in the left figure above.

Unified results saved to any folder in the Igor experiment. This is selected in the right figure above. Note, that in this case user needs to select folder which contains unified results as well as UnifiedFitIntensity_X, where X is number of the Unified results "generation" (remember, there may be many generations of results there). Note that this can be quite messy, if you do not know which generation to pick… The data for analysis are picked from the wave note of the selected wave.

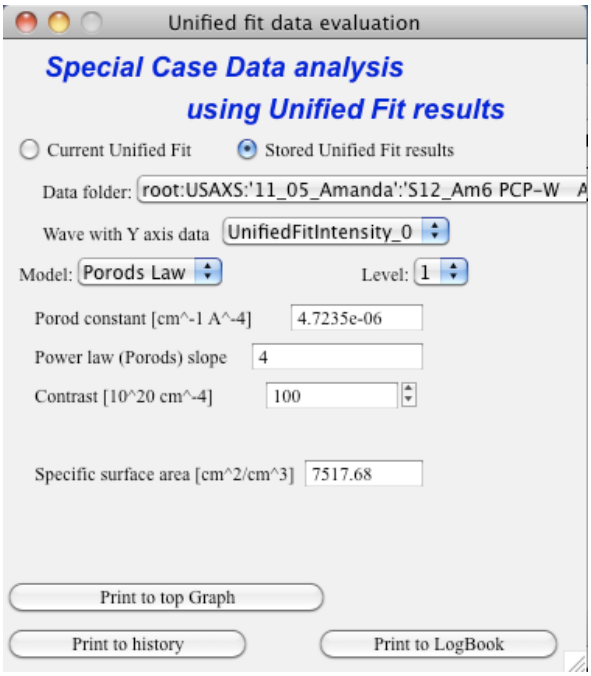
Then one can pick models:

Invariant



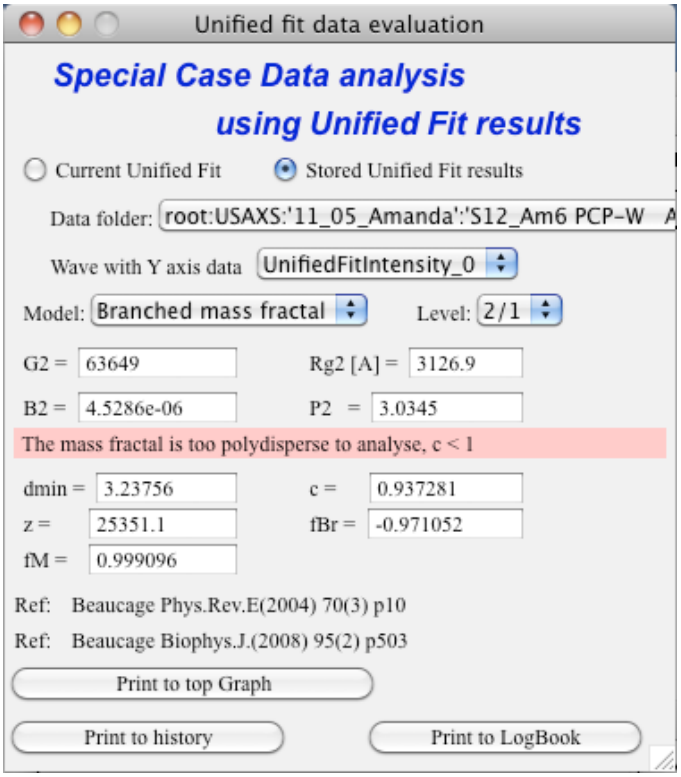
You need to pick also level for which to calculate the invariant. The invariant value is in the top field, user can input contrast value (if known) and if the data were absolutely calibrated and the contrast is known, the tool calculates the volume fraction of the phase.

Porods law



This will provide results ONLY, if the P for selected level is close to 4 (3.96 – 4.04). In that case, the tool provides Porod constant, P and calculates specific surface area – if the scattering contrast is provided. You need to have data absolutely calibrated.

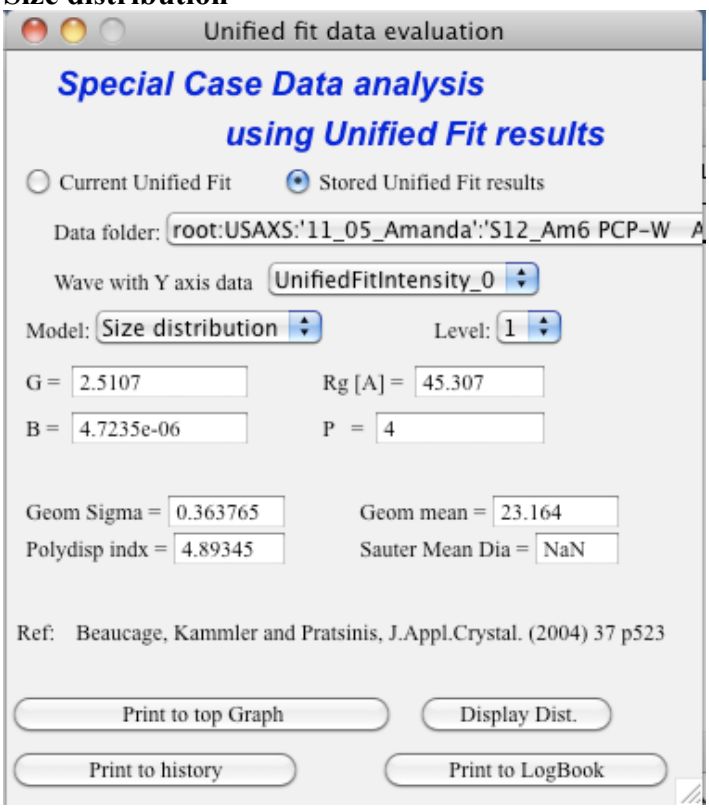
Branched mass fractal



Ok, this tool requires users to read the references. The code was provided by Greg Beaucage and provides results as expected. But I am not clear on what these numbers really mean. Any way, the references are on the panel itself.
Note, that when the calculations fail, the tool beeps and prints error message in the red box.

Note, to calculate all of the parameters, you need two levels – so there are choices like 2/1 (1 would be primary particles, 2 would be the mass fractal). But you can also calculate some parameters from only one level (dmin and c) and if you select only one level, parameters, which cannot be calculated, will be set to NaN.

Size distribution



In this case, parameters from one level can be used to calculate log-normal size distribution for the particles – which assumes the P is close to 4 (Porods law). The details are in the manuscript referenced on the panel. Please, read it.

Outputs of this tool:

User can get following outputs, using the buttons:
Print results to history area in Igor experiment. For example here is the results from the above Size distribution tool:

```
****************** Results for Size dsitribution analysis from Unified fit ***************************
    User Data Name : 'S12_Am6 PCP-W   A'
    Date/time : Analyzed using Unified Fit results from Sun, Feb 21, 2010  7:19:12 PM
    Folder name : root:USAXS:'11_05_Amanda':'S12_Am6 PCP-W   A':
    Intensity name : UnifiedFitIntensity_0
    Q vector name : UnifiedFitQvector_0
    Error name : ---

    Selected level : 1
```

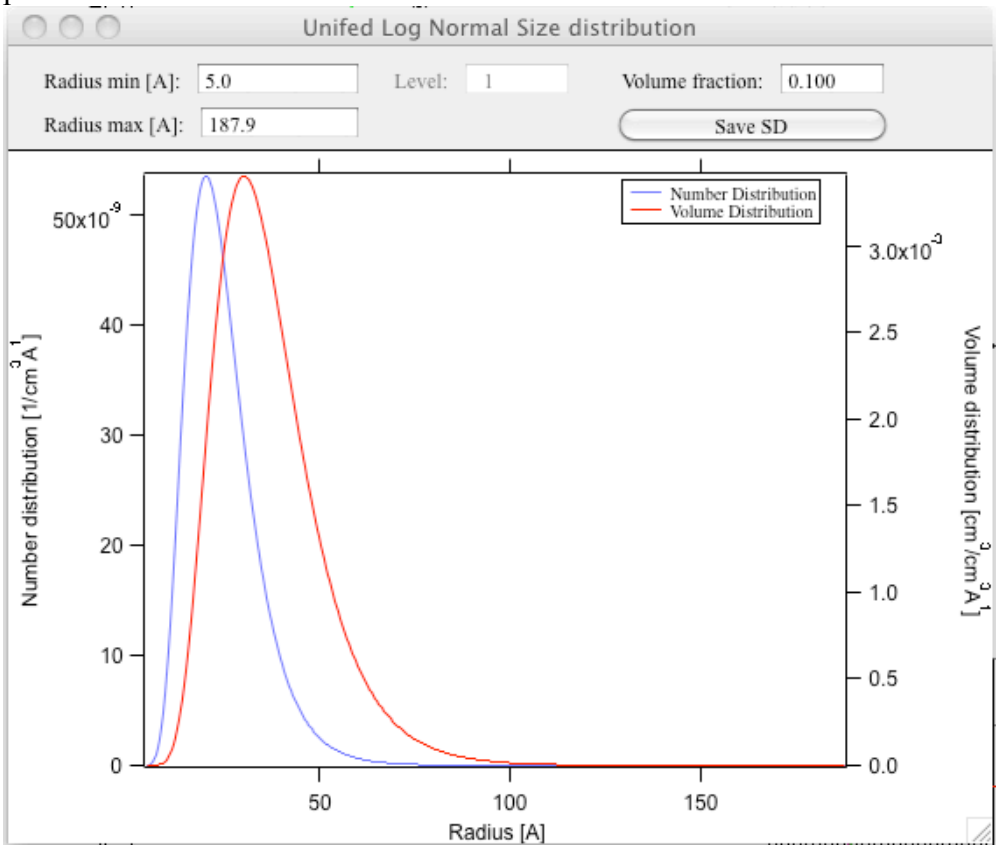
```
    G/Rg/B/P    2.5107          45.307     4.7235e-06          4
    Geom. sigma : 0.36376
    Geom mean : 23.164
    Polydispersity index : 4.8935
    Sauter mean diameter : NaN
    Reference : Beaucage, Kammler and Pratsinis, J.Appl.Crystal. (2004) 37 p523
**********************************************************************************************************
```

Print the same results into the Irena log. Remember, this is the log book (Igor "notebook" which manuy Irena tools save record of what is happening to.

Print textbox with the results to top (or Unified) graph. This is example of record for the Size distribution:

Size dsitribution analysis using Unified fit results
User Data Name : 'S12_Am6 PCP-W A'
Selected level : 1
G/Rg/B/P 2.5107 45.307 4.7235e-06 4
Geom. sigma : 0.36376
Geom mean : 23.164
Polydispersity index : 4.8935
Sauter mean diameter : NaN
Reference : Beaucage, Kammler and Pratsinis, J.Appl.Crystal. (2004) 37 p523

And for size distribution ONLY… Display distribution. This will calculate the log normal distribution for the parameters obtained.



User may need to change the Radius min and Max values (my code to guess these seems to fail miserably in some cases). The graphs shows which level was analyzed and enables user input of volume of the total volume of the particles in this size distribution. This is basically absolute scaling, as it looks like Greg Beaucage never worked out details of using absolute calibration of the data themselves. If this becomes important, I may be able

to develop the math myself. You can use for now for example value from invariant (which would be my choice for code anyway).

You can also save size distribution as waves for future euse (these waves are recognized as “results” in irena package now. You will get report in history area:

******
Saved Unified size analysis data to : root:USAXS:'11_05_Amanda':'S12_Am6 PCP-W A':
  waves :
UnifSizeDistRadius_1
UnifSizeDistVolumeDist_1
UnifSizeDistNumberDist_1

8. Modeling I (standard models)

This program allows user to create up to 5 populations of scatterers and combine small-angle scattering from them to compare with measured data. Fitting and optimization: Parameters can be least square fitted to measured data. From version 2.15 second optimization method is available – "Genetic optimization". See comments on Genetic optimization in the introduction.
Each population have it's own shape, distribution type, size, contrast, etc. Common selections only are which type of distribution probability is modeled – number distribution of volume distribution.

8.1 Details on Standard models mathematics:

What the macros do:

The macros model small angle scattering (SAS) using basic SAS formula:

$I(Q) = |\Delta\rho|^2 \int_0^\infty |F(Q,r)|^2 V^2(r) NP(r) dr$, where Δρ is contrast, F(Q,r) is scattering form factor, V(r) is the particle volume, N is the total number of scattering particles, Π(r) is the probability of occurrence of scatterer at size of r. This formula is, of course, replaced by summation formula with limited number of bins in radii. Therefore the formula coded in is following:

$$I(Q) = |\Delta\rho|^2 \sum_{r_{\min}}^{r_{\max}} |F(Q,r)|^2 V^2(r) N\Pi(r)\Delta r$$

This formula has been coded very many times… Following are comments, which address specific parts of this formula.

Summary of features:

1. Use of SAS data from Indra 2 data structure, QRS data structure – or any other data structure (need Q vector, intensity, error).
2. Allows creation of "model" only – does not require input data anymore (creates q values in user defined range)
3. Up to 5 customizable distributions of scatterers, each with it's own scattering contrast, particle shape, size distribution model, etc.
4. Three different distribution models available – Gauss (Normal), Log-Normal, and LSW (Lifshitz-Slyozov-Wagner used in precipitation theory involving Ostwald Ripening).
5. Automatic selection of radius distributions – ranges needed, bin widths etc. – with user selectable precision and number of steps.
6. Easy control of all parameters in one panel
7. Immediate comparison with the measured data in one graph in Intensity-vs-Q (log-log) plot and in Intensity*Q^4-vs-Q plot.
8. Independent graph for distributions of scatterers.
9. Fitting routine for parameters…
10. Automatic recording routine for parameters before and after fitting
11. Number of available form factors. Note, that it is relatively easy to add other shapes in the code, so if anyone needs (really needs) another shape, let me know… Note, that calculation speed of different form factors varies significantly depending on calculations needed to calculate involved integrals.

8.2 Interference

This code includes simple and very crude interaction calculation, which can be independently switched on or off for each population of the scatterers. **User should be aware of the crudeness of these calculations.**
The code used for calculations involves correcting intensity from a population of scatterers using this formula:

$$Intensity_{with\ interfernce}(Q, R) = Intensity_{without\ interference} * (1 + pack * SphereAmplitude(Q, Eta))$$

Where the *pack* and *Eta* are the two parameters of this model.
Note, that this is supposed to be valid for spheres. I am working on adding other methods for other arrangements of particles.
Remember: this method accounts in very crude way ONLY for interaction for particles in the particular population. If there are interactions among particles from different populations – which is very likely – these calculations have NO WAY to account for it.

8.3 Important Information

The code uses for all size related parameters Angstroems (10^{-10} m) or for Q vector (A^{-1}). In the case of scattering contrast, number distribution and any other volume contents centimeters (10^{-2} m).
This code uses everywhere Radius for scatterer size.

Distribution Ψ(r) and V(r)* Ψ(r) and distribution of r.

The code can work with distribution defined as for number distribution N*Ψ(r), where integral over Ψ(r) for all r is 1 and N is total number of scatterers or for volume distribution V_{tot}*Ψ(r), where integral over this term is equal total volume of scatterers. Internally, the code actually always works with number distributions (N*Ψ(r)), which, in the second case is calculated from the total volume of scatterers.

There are currently 3 different distributions built in the code, which can be used independently for any of up to 5 scatterers populations: Gauss (normal), Log-Normal, and LSW.
Gauss and Log-Normal distribution definitions were adopted from NIST engineering statistics handbook, www.itl.nist.gov/div898/handbook/eda . See http://www.itl.nist.gov/div898/handbook/eda/section3/eda366.htm
Standard complicated log-normal distribution is defined as follows (Allen, A.J., Krueger, S., Skandan, G., Long, G.G., Hahn, H., Kerch, H.M., Parker, J.C. and Ali, M.N. (1996). *J. Am. Ceram. Soc.* **79**, 1201-1212., Filliben, J.J. (2006). *Exploratory Data Analysis*, in *NIST/SEMATECH e-Handbook of Statistical Methods*, edited by C. Croarkin and P. Tobias, p. 1.3.6.6.9, available online at *http://www.itl.nist.gov/div898/handbook/*. Gaithersburg, MD: NIST.):

$$\underset{j=1,4}{\Psi_j(D)} = \frac{\phi_{j\mathrm{TOTAL}}}{\left\{2\pi \ln\left(\frac{D_{j\mathrm{MED}} - D_{j\mathrm{MIN}}}{D_{j\mathrm{MODE}} - D_{j\mathrm{MIN}}}\right)\right\}^{0.5}} \left(\frac{1}{D - D_{j\mathrm{MIN}}}\right) \exp\left\{\frac{-\left[\ln\left(\frac{D - D_{j\mathrm{MIN}}}{D_{j\mathrm{MED}} - D_{j\mathrm{MIN}}}\right)\right]^2}{2\ln\left(\frac{D_{j\mathrm{MED}} - D_{j\mathrm{MIN}}}{D_{j\mathrm{MODE}} - D_{j\mathrm{MIN}}}\right)}\right\}$$

The NIST definition is modified to be more elegant and parameters used by Irena package are as follows:
"Min" = Dmin
"Mean" = (Dmed - Dmin)
"Sdev" = sigma = ln((Dmed-Dmin)/(Dmode-Dmin))

The LSW distribution has been accepted from a source by J. Nasser, A. K. Kuruvilla, and J. E. Smith Jr. These authors in their manuscript on the web (www.space.gc.ca/science/space_science/paper_reports/spacebound97/materials_sciece/….) refer to distribution by Lifshitz, Slyozlov, and Wagner:

$$\Psi(r) = \frac{81}{2^{\frac{5}{3}}} \frac{\rho^2 \exp(-\frac{\rho}{1.5-\rho})}{(1.5-\rho)^2(3+\rho)^{\frac{7}{3}}}, \rho < 1.5$$

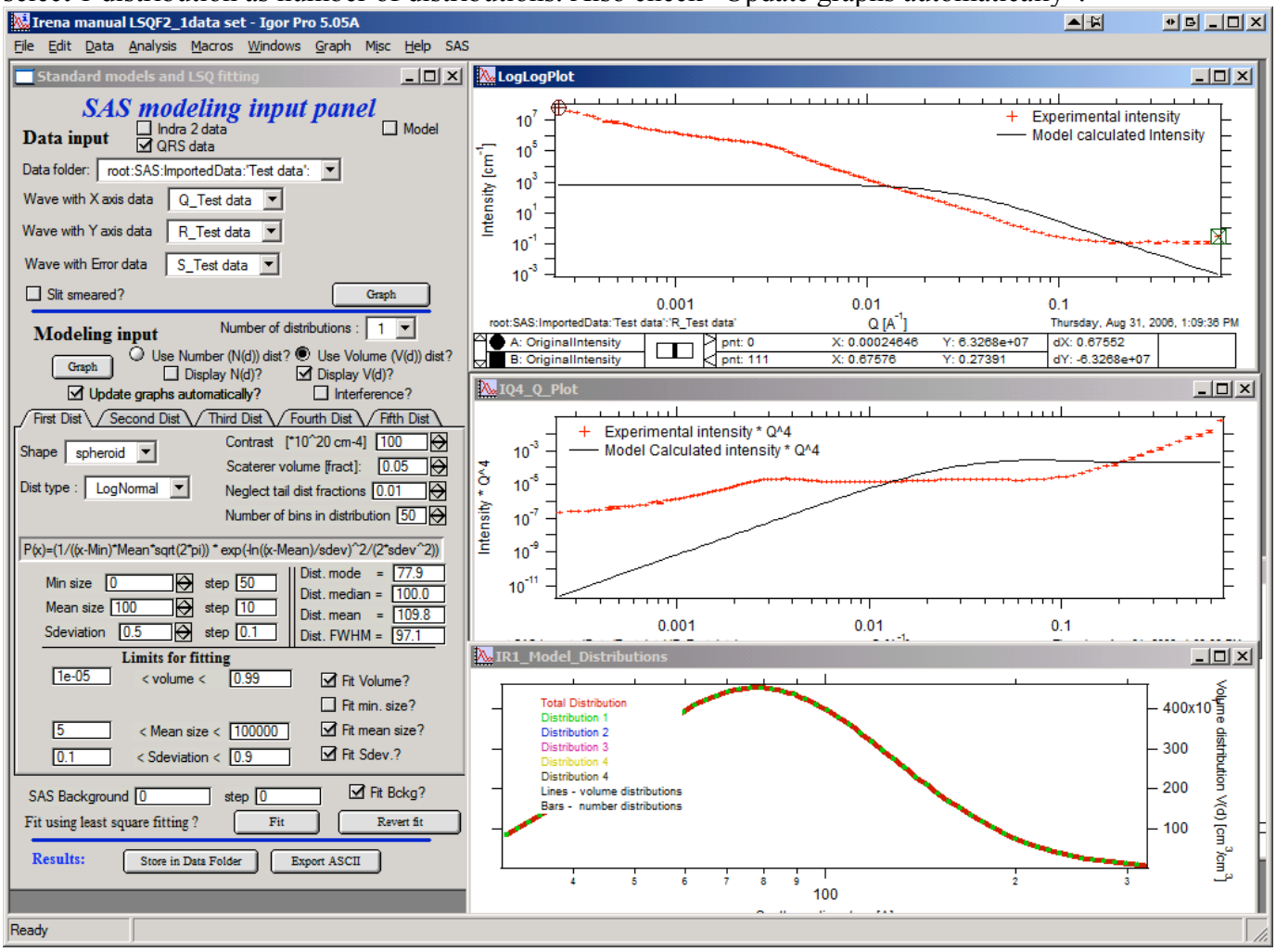
This is the particle size distribution predicted by LSW in their theory of Ostwald Ripening.

Each distribution in this type of problems needs an appropriate selection of radial bins. Appropriate selection is actually problem – too many bins cause too long calculation times, narrow range of radii causes some significant volume of scatterers to be neglected, etc. In this code I take a different approach, which is important to explain properly:
For each distribution I create cumulative distribution (if exists using formula, if not numerically). Using user input value I select range of radii in which the value for cumulative distribution is between this value and (1-this value). This causes, that only the tails, for which the cumulative probability is below the user selected value are neglected, giving user full control of the precision in which we/she wants to model the data. Then radial bins are calculated, so their spacing for cumulative probability is the same. This causes that the bins have varying width – are narrowest around the areas, where probability function changes fast and wider in the tails. This should provide the best possible method for using the binning method, I hope…
All of the code handles bins of varying width…

F(Q,r) problem – applicable ONLY to integrated spheroid

For the case of **integrated spheroids ONLY** - *rarely addressed problem is related to usual method of calculation of F(Q,r), independent of selected particle shape. In usual method of modeling – using bins this problem is usually neglected. Standard method is to take for radius the center point of the bin, and calculate F(Q,r) for this point. However, this may be very incorrect - the F(Q,r) is a strong function of Q*r (with period of pi). Through the Q range and size range studied, the number of periods in pi within the bin width * Q varies strongly. Taking just center of the bin for calculating F(Q,r) results in nearly random selection of the r for this calculation and can result in significant error. Calculated value may be very far from average F(Q,r) value, which we should properly used.*
In case of data from USAXS instrument we at least have no problem with definition of Q – the Q resolution is very high, otherwise we would have to worry about the Q variation within the Q point - smearing…
Anyway, to avoid problems with the oscillatory behavior of the F(Qr) the code takes at least 3 – and maximum 61 F(Qr) values within the bin in radius (at least start, middle and end of the radius bin point), linearly distributed in the bin, multiply them by appropriate V(r) and then average the result. The number of points within the bin is obtained as floor(3+abs((10(QRMax-QRMin)/pi))), with maximum being 61.*
This causes significant increase in the calculation time… Keep this in mind.
--- end of part valid ONLY for spheroids
The above does not apply for other shapes – globs by definition do not exhibit this problem and I have not included this complication for other shapes. The standard spheroid also do not have this included – if you want to use this integration method, use even for spheres "integrated spheroid" and aspect ratio 1.

Select “Modeling I” from “SAS” menu. This brings up control panel. Select available data set, “Graph”, and select 1 distribution as number of distributions. Also check “Update graphs automatically”:

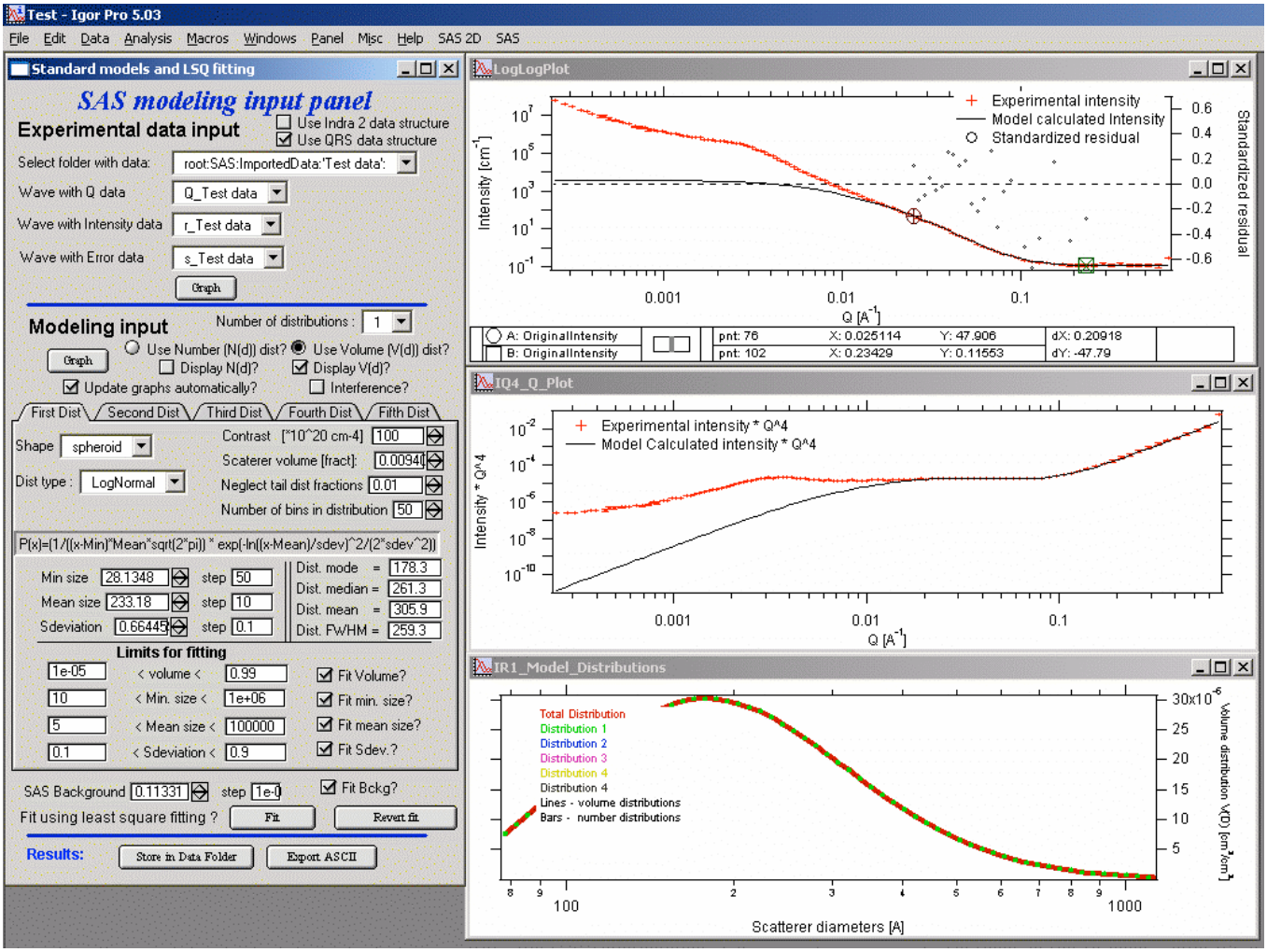


Top graph is log-log plot of Intensity vs Q, lower plot is Intensity * Q^4 vs Qvector (=S(Q) as known from scattering theory).
This is default distribution for distribution 1. Shape is spheres (spheroid with aspect ratio 1).

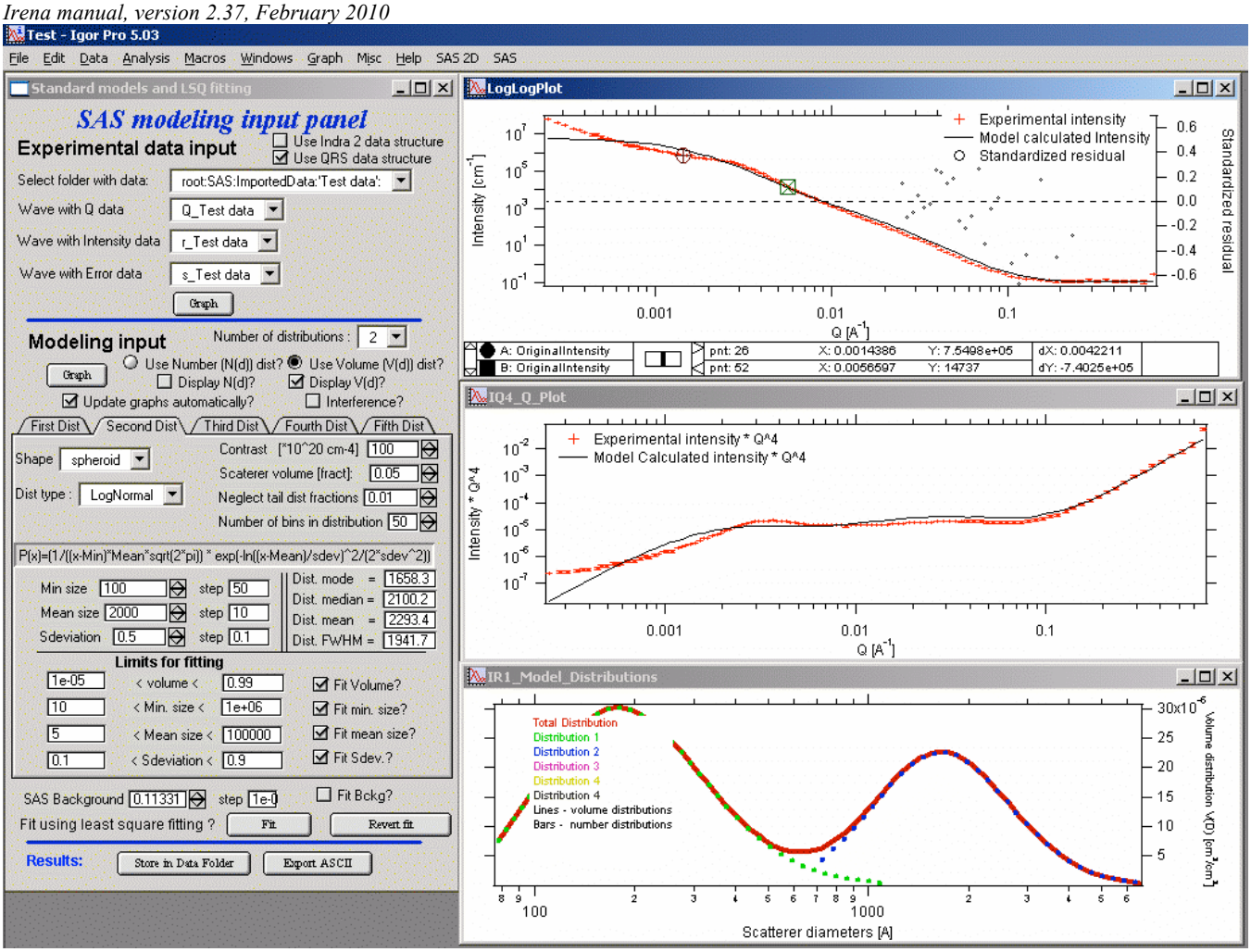
Distribution type is Log-normal (other options: Gaussien and LSW).
If the distribution parameters seem unusual, the formula used is posted above the parameters. Values for contrast, volume fraction of scatterers (0.05 = 5%) and number of bins in distribution are quite clear in their meaning.
Parameter “Neglect tail dist fraction” is more complex… When I create the distribution according to parameters user provides, firs I find the middle of the distribution (maximum). To create cumulative distribution and use that to distribute bins so the change in cumulative distribution in each bin is the same. However, there is need to terminate the binning at some minimum and maximum size. To do so I need to know, what amount of cumulative distribution user is willing to neglect. So if this parameter is set to 0.1, then binning in radii (diameters) is done so there is less than 1% probability for radii (diameters) smaller than smallest bin size and less than 1% probability that it is larger than largest bin size. This is necessary, since the Gaussian and Log-Normal distributions used here are unlimited.

Parameters Dist mode, median, mean and FWHM (full width at half max) are calculated numerically for user to provide some sensible information on the probability distributions. They cannot be changed.

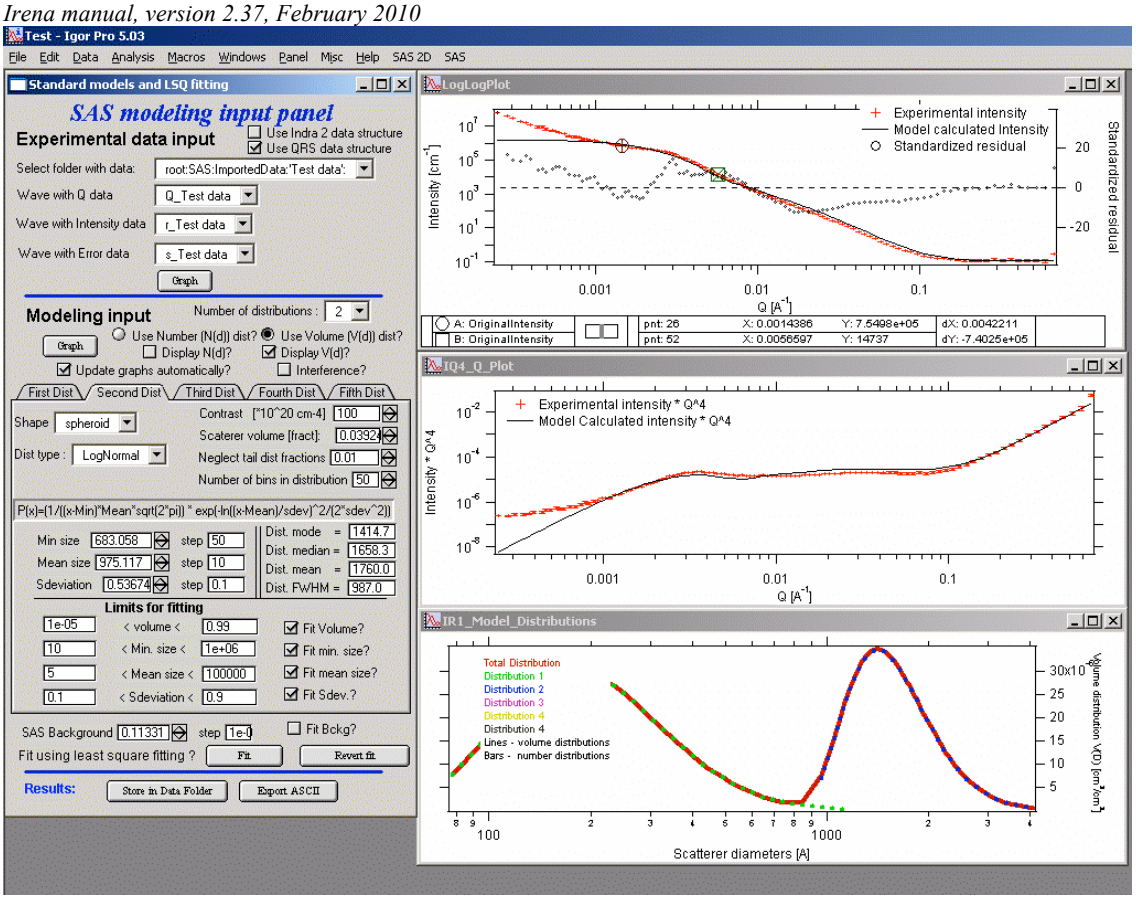
Parameters used in the distribution formula can be modified and fitted. Modify them to fit the Guinier area of the smaller particles. Note solution below, which is not bad starting point. Note input range by selecting data with cursors in the TOP graph (log intensity vs log Q). Now select area between points 69 and 100 (only top graph can be used for input!!!) and check Fit Volume, Fit min. size, Fit Mean size, Fit Background checkboxes (Fit Sdeviation causes significant instabilities, so use sparingly). Check limits (FitVolume lower limit may need to be decreased, for example).



Now, let's add second population of particles. Select number of distributions 2, and click on the tab with "Second Dist" name. Increase the Mean size for this population to 2000A and possibly modify other parameters according to following parameters.

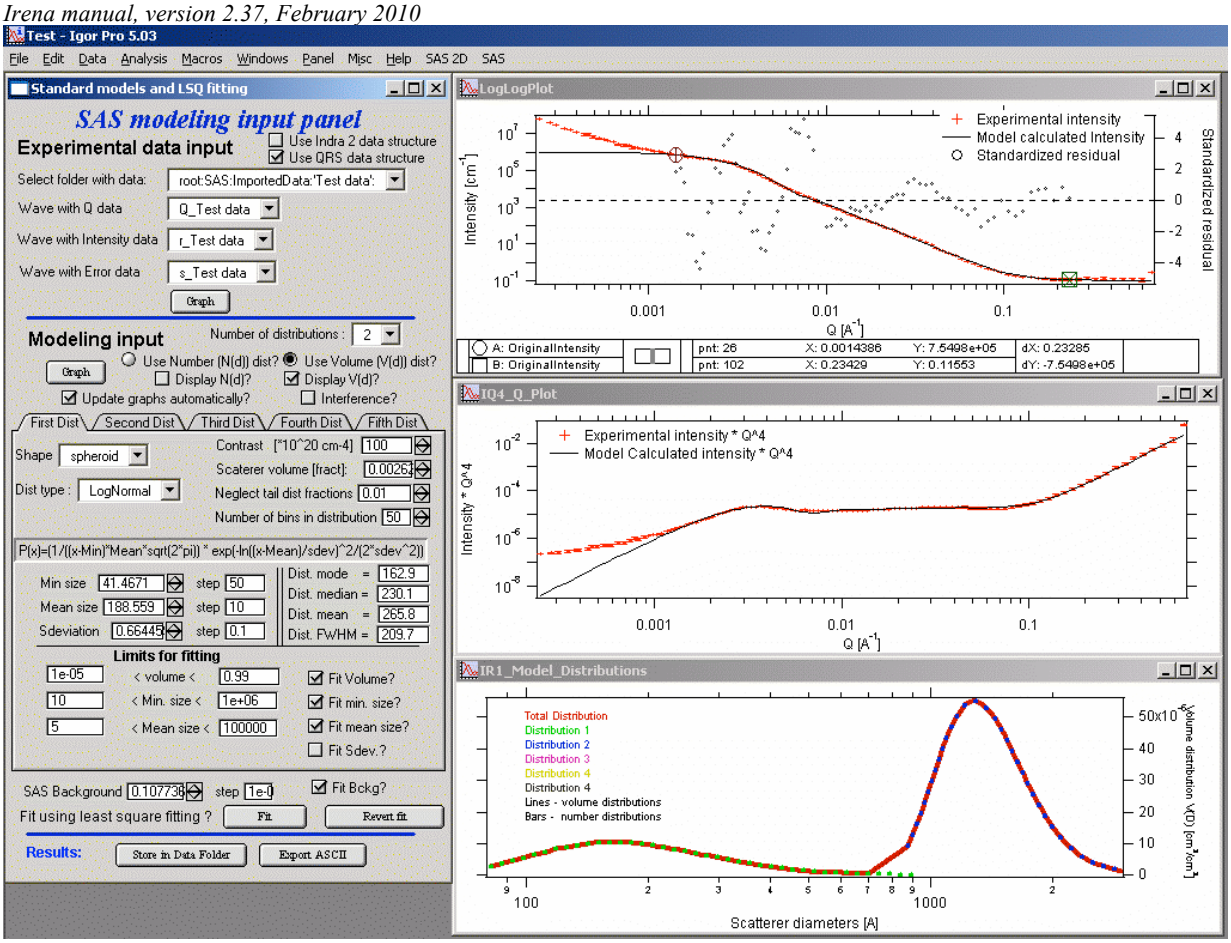


This is not bad guess. Note, that as the second population became larger, if you display number distribution the distribution line nearly disappears from the bottom graph. There are only few of large particles there. Therefore in this case is better to display volume distribution, which will show the second population. However, choice of user if to model volume or number distribution needs to be made based more on physical basis…
Select point 29 to 62 in the top graph using cursors. In the Tab of the First distribution uncheck all fitting boxes, uncheck Fit Background checkbox and in the tab for second distribution check Fit Volume, Fit Location, Fit Scale checkboxes… The fit.



This is reasonably good local fit for second population. Note the straight line between the distribution for the first population and second population – this is where there are no datapoints, since first population already ended and second did not start. This area will get smaller if the "Neglect tail dist fraction" is set smaller. However, if this number is very small, or distributions are close to small sizes, problem with bins with negative diameters appears. The code just truncates the binning at 2A or so, so there can be artifacts for distributions which have significant volume at smaller scatterers…

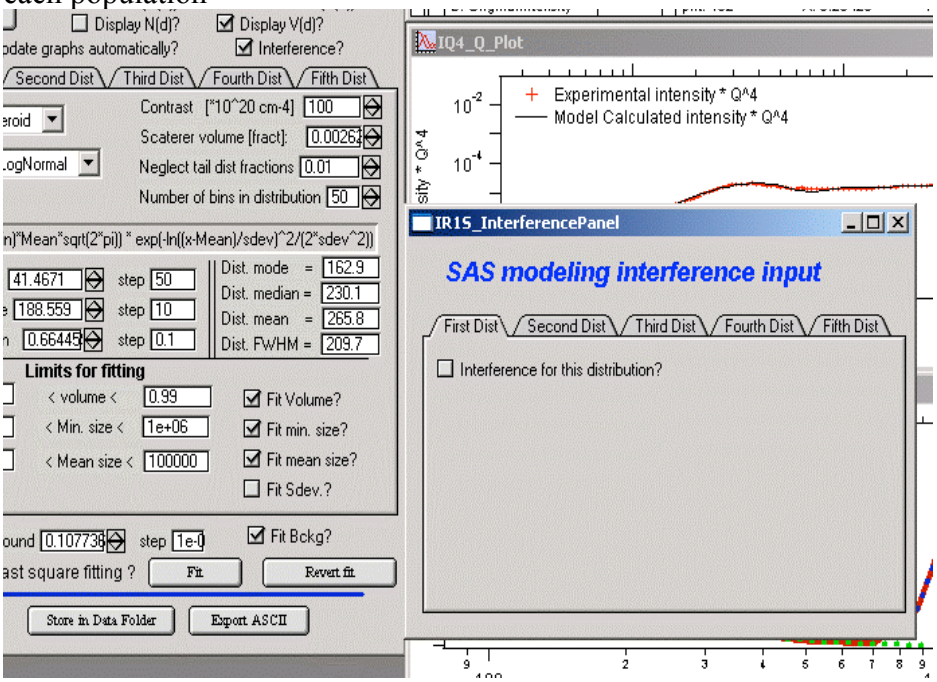
Now we can optimize all parameters at the same time. Select point about 29 to 102 using cursors. Check major parameters from First distribution and fit all together.
Final fit may look like:



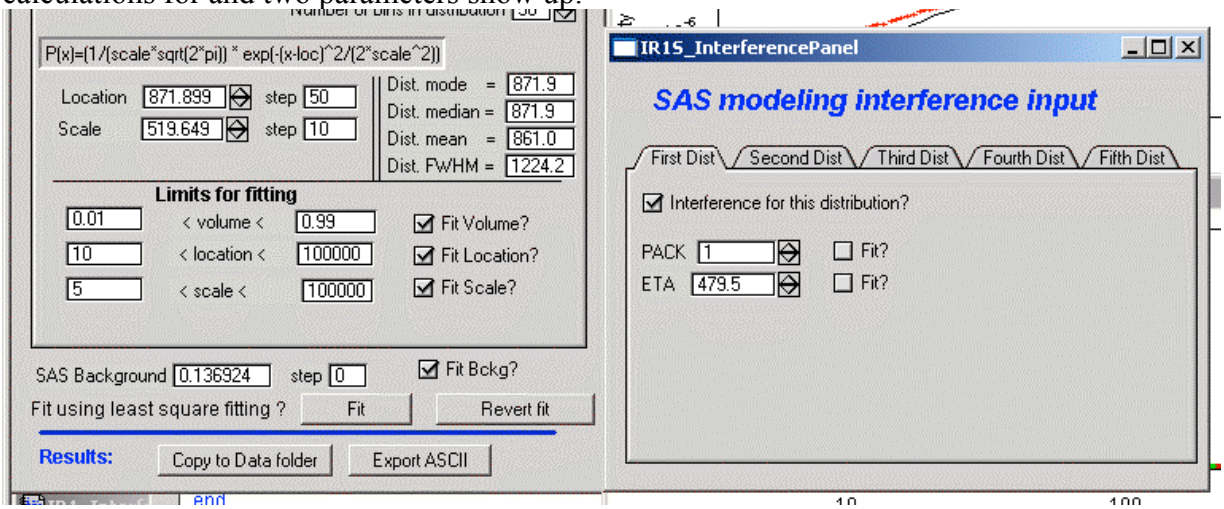
Note, that fitting the small-Q powerlaw with this model is not easy, since there is no real size information about those scatterers. Therefore I will not attempt to do it at this time.

Adding interactions

Check the checkbox “Interactions” on the control panel. New panel pop-ups with Interference control tabs for each population



Check the checkbox “Interference for this population” for population, you want to include interference calculations for and two parameters show up.



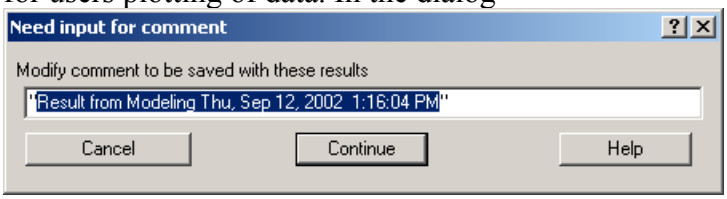
These parameters can be fitted by checking “Fit?” checkbox, when other necessary fields appear. Note, that in my experience the fits are not very stable. Need to have really good estimates for these parameters. For more on these parameters, look into the Unified fit section, since this is the same code used here as in Unified.

Saving results

It is now possible to save data back to data folder where the measured data came from or export ASCII file.

Copy to data folder

User can copy Igor waves (as in above methods) with comments, which can be recovered in the future or used for users plotting of data. In the dialog



user can fill in any text description (in the quotes).
Following waves are created in the data folder. _0 is index of users solution saved in the folder, so user can have multiple solutions. These solution can be recovered in the future as in previous cases…

ModelingDiameters_0	diameters for the distributions
ModelingNumberDistribution_0	number distribution
ModelingVolumeDistribution_0	volume distribution
ModelingIntensity_0	intensity from the model
ModelingQvector_0	Q vector for the model

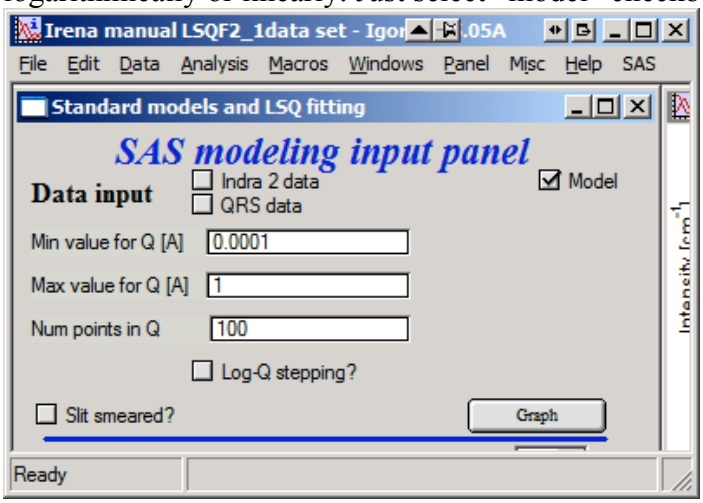
Note, that each of these waves has again wave note with all parameters from this model.

Export to ASCII

Exports ASCII file with all parameters (wavenote from solution waves above) and columns of data with the distributions and intensity-Qvector.

Modeling only

In order to facilitate modeling only of SAS from population(s) of scatterers, user does not need any input data. These data will be created – in Q range user selects, with number of points user defines and spread either logarithmically or linearly. Just select "model" checkbox at the top of the Data Input part of main tool panel.



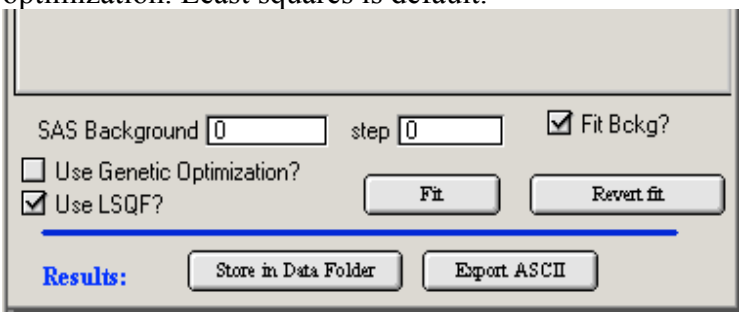
Select proper values and log-Q stepping, then push "graph". Tool will function with no input data now…

8.4 Fitting

General comments

This code uses either Igor built in least-squares fitting routine or Genetic optimization.
Least square fitting: The SAS data (intensity, Q and error) always extend over large ranges – often many decades – this routine is not the best one for parameters optimization. Therefore it is imperative to have a good starting guess for the parameters, properly select range of values, the fit is allowed to look for solution and fit limited number of parameters at a time.
Genetic optimization: very good for cases when number of local minima can make least square fitting unusable.

Note change on the panel: Lower left corner – the two checkboxes switch between least squares and genetic optimization. Least squares is default.



For genetic optimization user is presented with review of parameters which will be fitted and review of range which will be probed. It is imperative to have limited and sensible range of parameters to be fitted!!!

8.5 Logging feature

This feature is not finished and works only for standard models. It is planned update for future to make sure these records are useful…
User can see the notebook by selecting second item in the SAS menu "Show SAS logbook". This area is at this time under development, but this is current status of what is written in this logbook:

This is log results of SAS fitting with modeling macros Irena.

1/5/02, 5:47 PM

****
****
****
Parameters before starting Fitting on the data from: root:USAXS:'S5_Al2O3 1um':

Number of modelled distributions: 1
SAS background = 0.15, was fitted? = 0 (yes=1/no=0)
********** Distribution 1
Particle shape: sphere
Distribution type: LogNormal
Contrast 120
Volume 0.09 , fitted? = 0
Location 250 , fitted? = 1
Scale 300.1 , fitted? = 1
Shape 0.5 , fitted? = 0
Mean 575.21
Median 550.12
Mode 483.83
FWHM 291.36

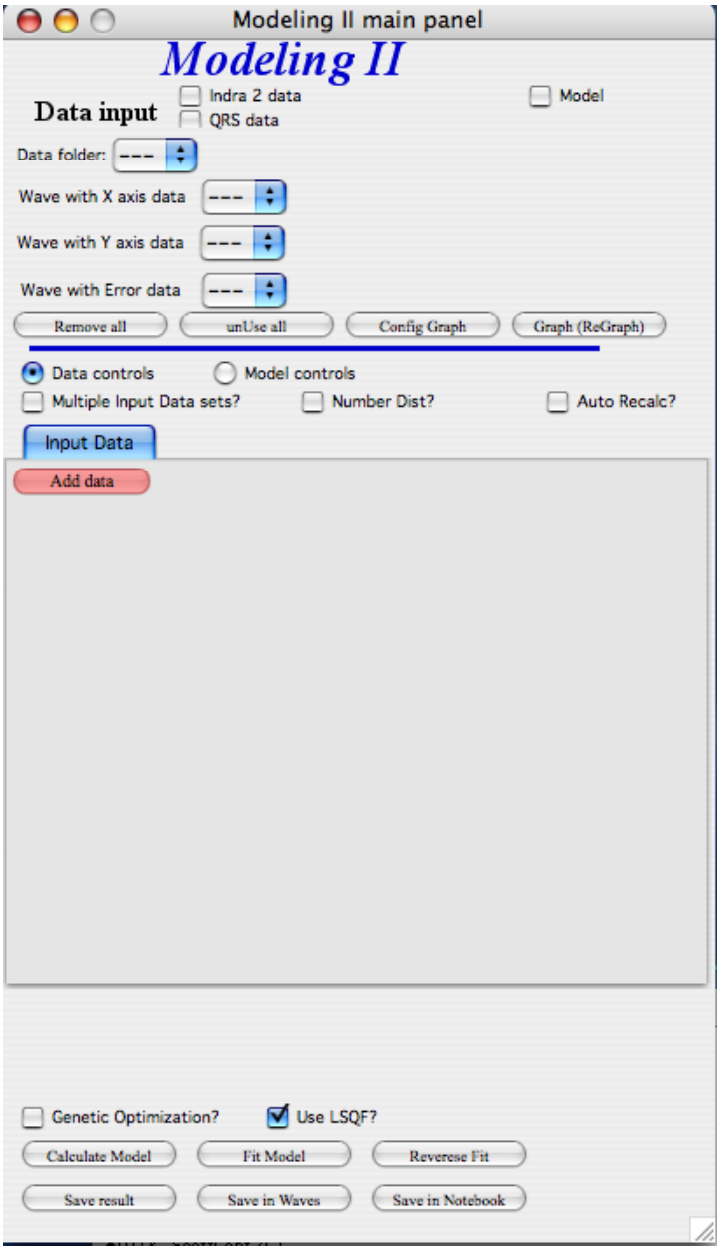
****
Results of the Fitting on the data from: root:USAXS:'S5_Al2O3 1um':

Number of fitted distributions: 1
Fitting results:
SAS background = 0.15, was fitted? = 0 (yes=1/no=0)
********** Distribution 1
Particle shape: sphere
Distribution type: LogNormal
Contrast 120
Volume 0.09 , fitted? = 0
Location 278.88 , fitted? = 1
Scale 278.32 , fitted? = 1
Shape 0.5 , fitted? = 0
Mean 575.21
Median 550.12
Mode 483.83
FWHM 291.36

Fit has been reached with following parameters
Chi-Squared 910.31
Points selected for fitting 28 to 109

The record will get significantly more complex in the future and, if using more populations of scatterers, also longer…

9. Modeling II



9.1 Use of this tool

This tool is : ***advanced; complicated; challenging to use – but really powerful***.

The use of this tool is recommended to experienced users or in cases, when the simpler tool from chapter 9 is not sufficient.

This tool will do all the LSQF ver.1 will do and much more. There is penalty for the much more…

Features:

1. Multiple input data sets at the same time. Up to 10 input data sets can be loaded at the same time and fitting/modeling can be done to one or more data sets at any time.
2. Up to 6 populations of scatterers can be combined. Any population can be arbitrarily "switched on and off".

3. Each population can have different contrast for each data set. Note, that this means up to 60 contrasts potentially.
4. Form factor parameters can be fitted. That does not mean the data must allow them to be fitted, but the tool will allow any form factor parameter to be fitted.
5. Structure Factors – library of 5 structure factors to be used. See Structure factor description in the "Form factor and Structure factor description" pdf file.
6. Optimized for speed - much faster fitting when "semi-auto" R distribution is selected for complicated form factors.
7. Input of "model" data when no real data are loaded, but only q-data are created to model SAS from microstructure without the need for measured data.

Logging during use of this tool – This tool similarly to other Irena tools logs state of the parameters before and after each fitting. The purpose is to enable at least some recovery, if fitting goes to never-never land and user accidentally looses the "recovery" option by using revert fit (by hitting the fit button second time etc…). The notebook can be pulled up by using menu SAS -> Other tools -> Show SAS Logbook. This logbook is mess of information, but all necessary information to recover should be there. It is not formatted in any particularly meaningful way, though. As output use the "Save in notebook" option discussed later in this chapter.

9.2 Theory behind this tool

Actually, theory is the same as for LSQF from chapter 9 – up to 6 populations of scatterers can be created, the SAS from them can be calculated in dilute limit and summed together. Approximation of interparticle interactions (Structure factor) can be applied independently to any of the populations. For comments on used size distribution shapes (log-normal, LSW) please refer to chapter 9.3. It is important and useful to understand.

9.3 Use of this tool for SINGLE input data set.

Start tool from SAS menu. At this time it is titled "Modeling II".

GUI description

Note in following image, that the tool has somehow different GUI. The lower 2/3 of panel change depending on selection of radio-buttons "Data controls" and "Model controls". See figure below. Also, if only one data set will be used, make sure the checkbox below "Data controls" is unchecked (only one tab "Input Data" is visible).

Note, that there are few buttons just under the Data input popup controls: "**Remove all**", "**unUseAll**" "**Config Graph**" and "**Graph(reGraph)**". These are tools to control global behavior or resent the tool.

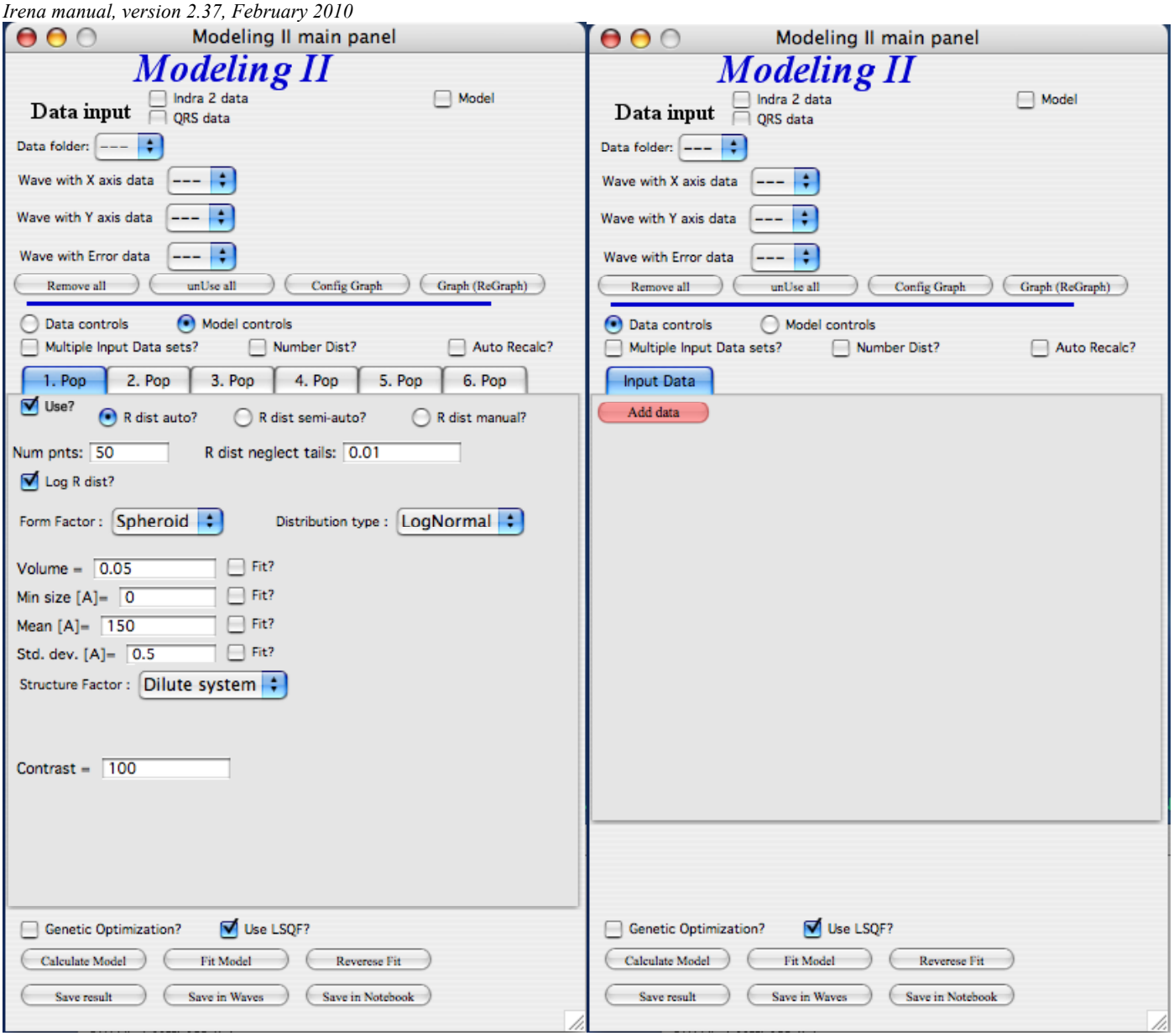
"**Remove all**" button removes all input data from the tool.

"**unUseAll**" button sets all of the input data sets to not to be used. (useful only with multiple data input)

"**Config Graph**" opens control screen for graph controls (font size etc.). These values are common for all tools (once I propagate them through whole package).

"**Graph (reGraph)**" button creates the graph or forces redraw of the graph.

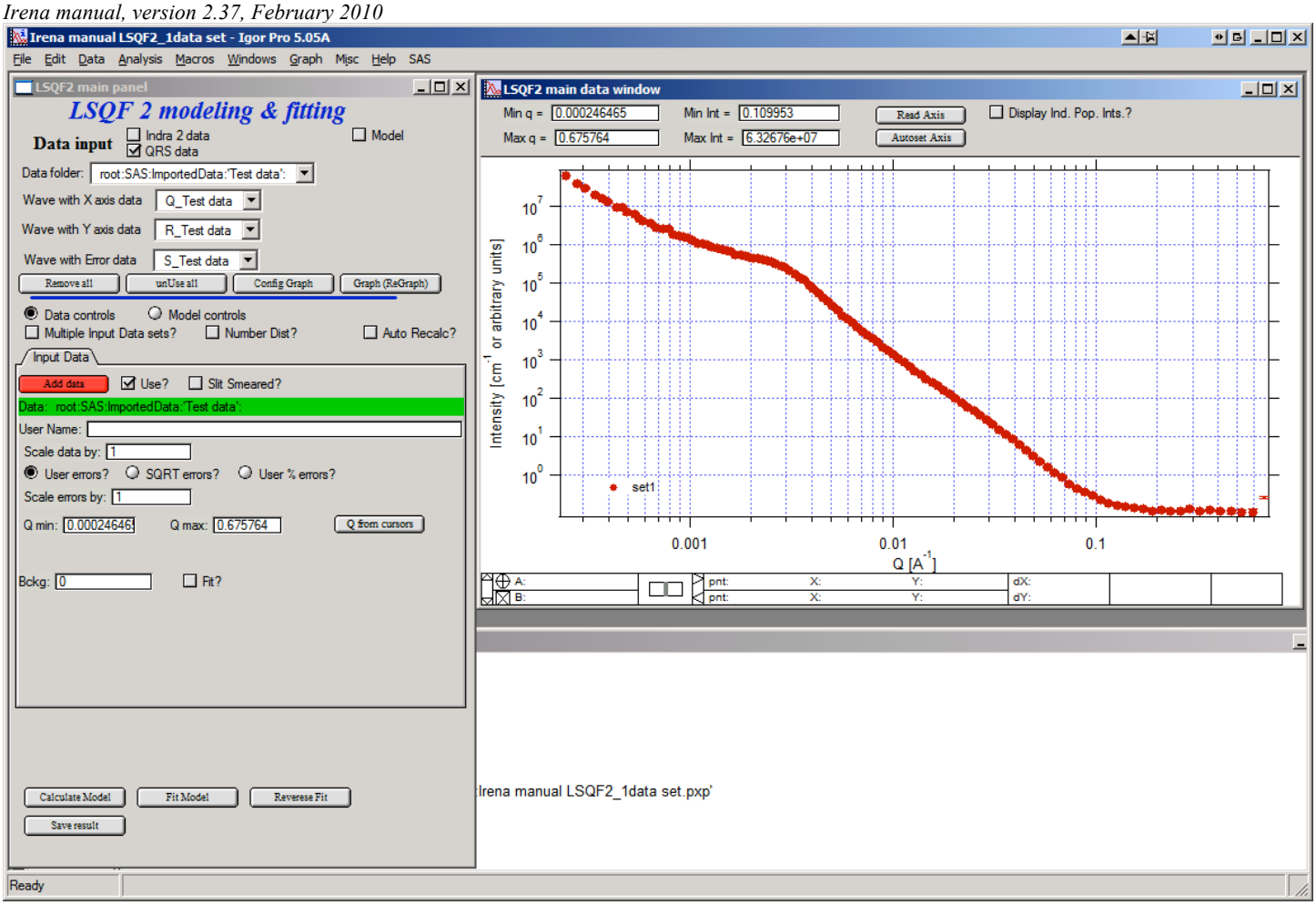
Note one more checkbox which is worth mentioning here… It is little bit lower, on the right hand side and is called "Auto recalc?". If checked the model will be recalculated with every change of any parameter (except Form factor parameters, which cannot trigger this). Use only on fast computers and simple enough model, or it can be tedious..



Data controls

The data available in the test.pxp file distributed with the irena package are in qrs structure, so select "QRS data" and pick the 'Test data".

To load data into the tool use the red button "Add data" on the left top corner of the Input Data tab.



Description of parameters on the Input data tab:
Checkbox "**Use?"** allows to select if this data set is used in the tool. This is really useful when multiple data sets are used.
Checkbox "**Slit smeared**" if slit smeared data re used, select. Note, that if checked field for slit length will appear.
"Data" field. This field contains path to data within Igor experiment. Cannot be edited.
"User name" user editable name for the data. Will be used in the graph – needed to make sense in case of use of multiple input data. If empty, default name will be used (not very informative).
"Scale data by" field – user can scale data here. For example some data may need to recalibrated, converted to 1/cm or whatever. Ideally should not need to be used.
Radio buttons "**User errors"**, "**SQRT errors"**, and "**User % errors"** – what type of errors to use for this particular data set? User errors are provided by wave with error data, SQRt errors are square root of intensity and when % error is used, the error is set to 1 % of intensity.
Scale errors by:" allows scaling errors by factor. Errors are produced using method selected above and then scaled by the factor user provides here.

"**Qmin**" and "**Qmax**" – selection on fitting interval of data – can be typed in or using the button "**Q from cursors"** can be read from cursor position. Only data within this interval will be used for fitting.

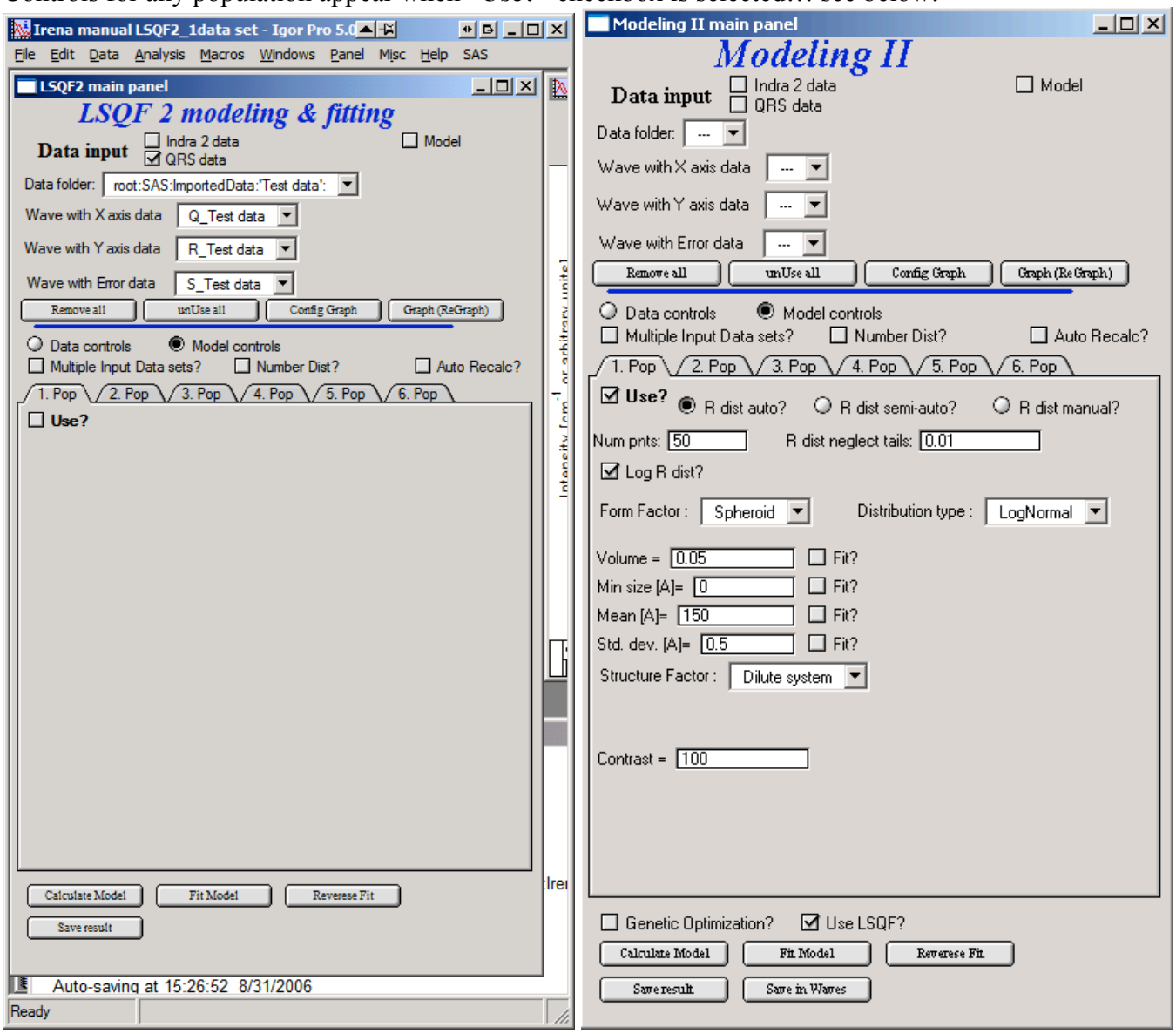
"Bckg" Background for this data set. Can be fitted ("**Fit?**" checkbox will open fields for Min and Max limits for fitting…

Further controls are likely going to appear…

Model controls

Model controls become available by selecting "**Model controls"** radio button above the tabs.
Note the checkbox "**Number Dist?"** – if checked the distribution will be considered to be number distribution, if unchecked (default) the distribution is volume distribution.

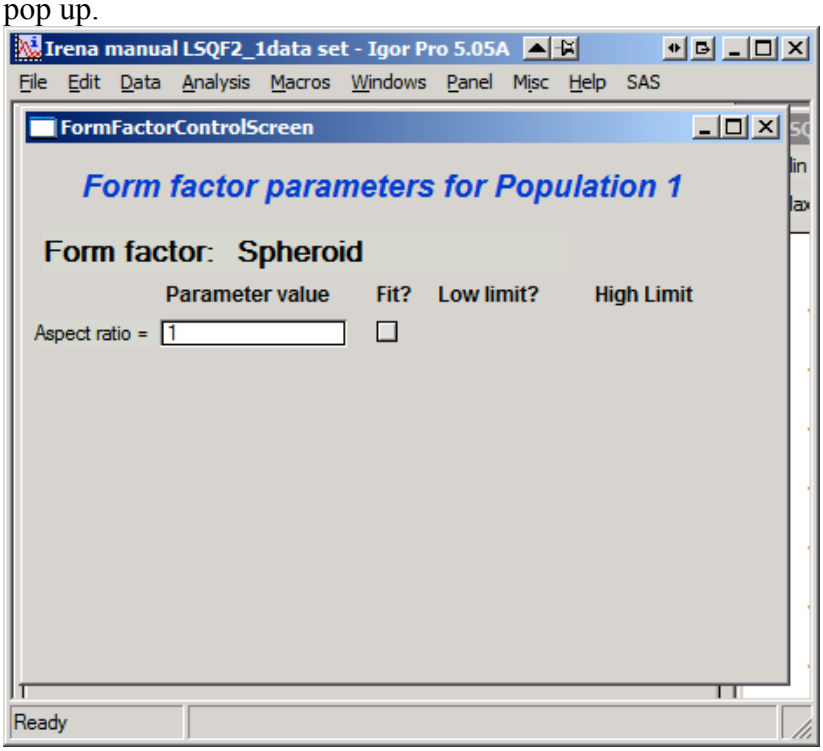
Controls for any population appear when "**Use?**" checkbox is selected… see below:



Description of controls:
"**R dist auto?**" distribution of radii selected automatically for given distribution. As in the older LSQF (chapter 9) the R distribution here is selected in such way, that densest points in R are at the middle of the distribution (around maximum) and then they spread with large and larger steps.
"**R dist semi-auto**" same as above, except the R distribution is not being changed during fitting. Therefore one needs to be close to final solution when starting fitting. But this way the fitting can be MUCH faster for complicated form factors. Since the R points and q-points do not change during fitting, G matrix (which is cashed internally for each population and data set) is calculated ONLY once. Major time saver…
"**R dist manual?**" Manually input min/max R for each distribution. Opens control fields needed for input.
"**Num pnts"** Number of points in R distribution. Use sensible numbers. Large numbers will take a lot of time.

"**R dist neglect tails"** same meaning as in LSQF (chapter 9). Basically what fraction of volume of size distribution can be neglected. Allows truncation at small/large sizes - defines Rmin and Rmax for automatic/semi-automatic R distribution method.
"**Log R dist?"** – select to have R points logarithmically distributed. If unchecked, linearly distributed bins in R will be created.

"**Form Factor"** – select form factor from list of available form factors. May open another control screen for parameters of the form factors. To get this controls screen again, re-select the form factor and the screen will pop up.



This is example of screen for Spheroid. Note, that there is one parameter for this Form factor (aspect ratio). This parameter can be fitted in this tool. By selecting "**Fit?**" checkbox, low an high limits fields will appear.

"**Distribution type**" select "lognormal", "Gauss" or "LSW". Definitions are in LSQF (chapter 9). Parameters for these distributions are now separate, so one can go among them and the parameters will not be reused/lost from previous use of that particular distribution type…
"Volume" – volume of scatterers in this population. "**Fit?"** checkbox allows fitting. Fields for min/maxc values will appear. When volume is changed manually by typing in this field, min and max are automatically set to 1/5 and 5x the typed value. Therefore, it is important to first set the value and then, if necessary change the limits. Not the other way around!!!

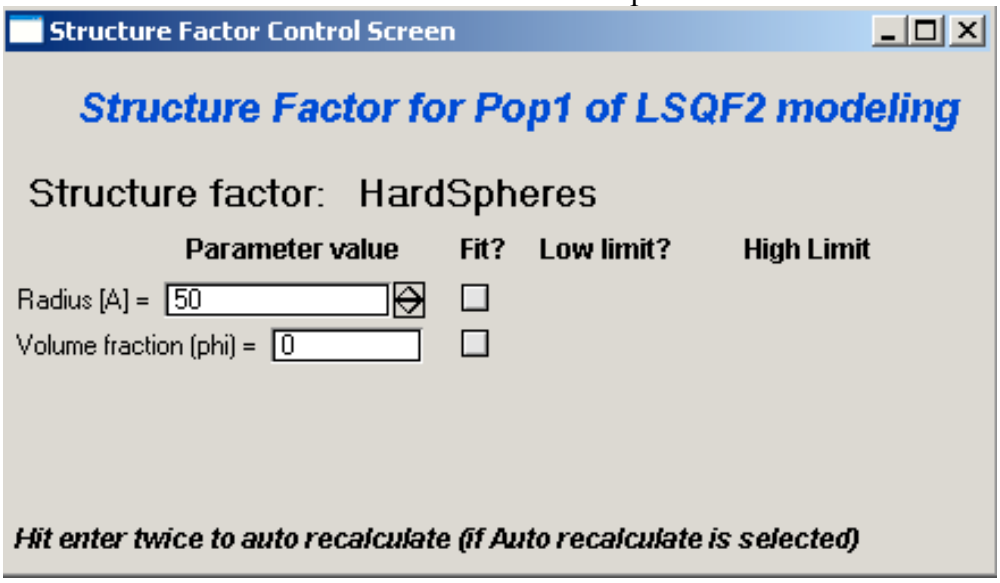
LogNormal parameters
See details in the chapter 9.3 for details… Formula:
P (x) = exp(-1*(ln((x-MinSize) / meanSize))^2 / (2*SDeviation^2)) / (SDeviation*sqrt(2*pi)*(x-MinSize))

"**Min size",** "**Mean"** , "**Std Dev."**

Gauss
Has just two parameters: "**Mean size"** and "**Width"**.

LSW
Just one parameter: "**Position"**. For details see chapter 9.3.

"Structure factor" Popup allows selection of one of included structure factors (see pdf file with description). The structure factors have their own screens and parameters can be fitted. See below for case example:



Note, that due to quirk in Igor way of controlling updates you need to hit enter twice to automatically recalculate the curve (when checkbox on main panel is selected)…

"Contrast" field – input contrast. Only one contrast in case of single input data set.

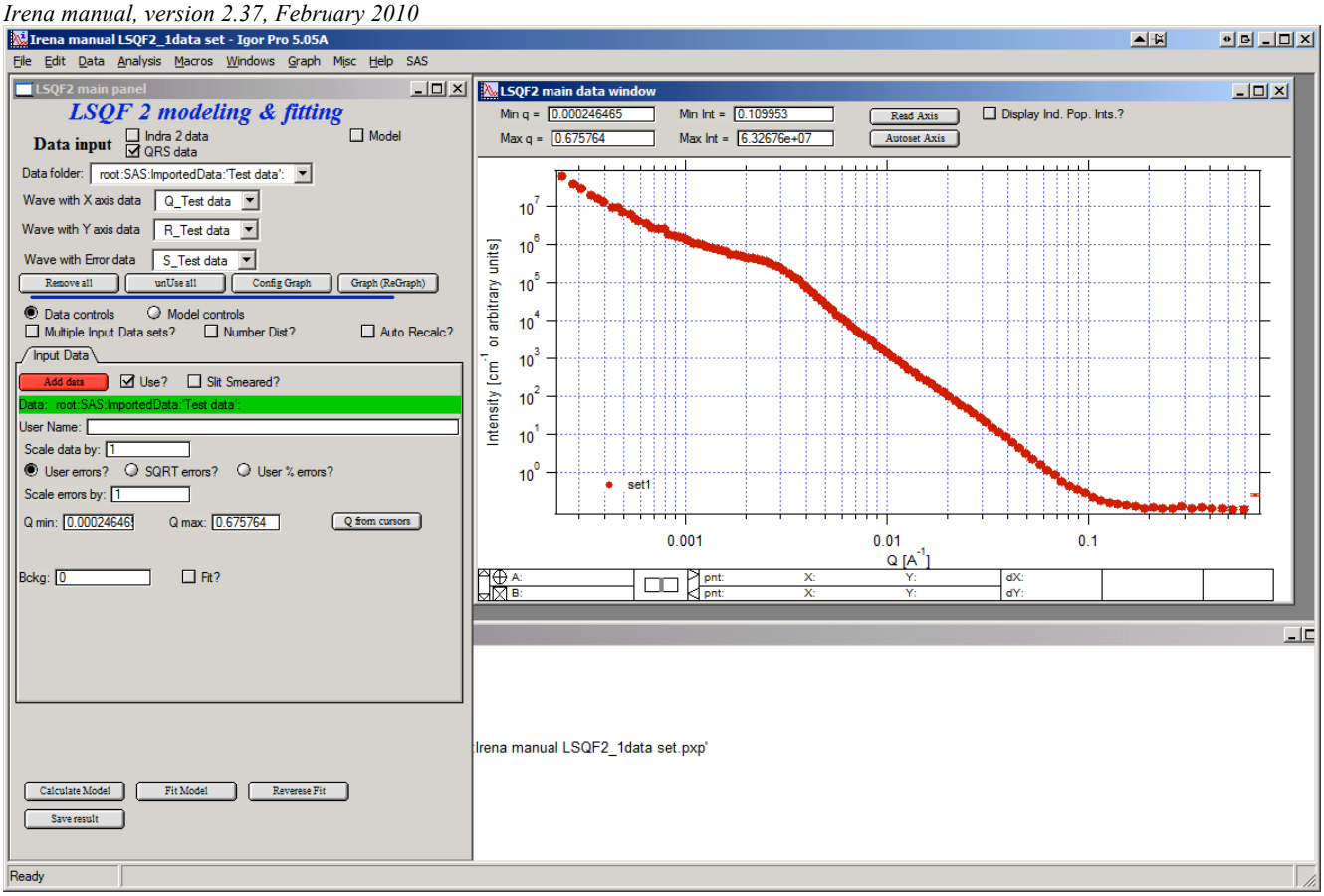
Last few buttons

Under the tab area there are few more control butons.
"**Calculate model"** calculates Intensity for current model.
"**Fit model"** Runs fitting with currently selected parameters to fit.
"**Reverse fit"** recover parameters stored before the current fit run.

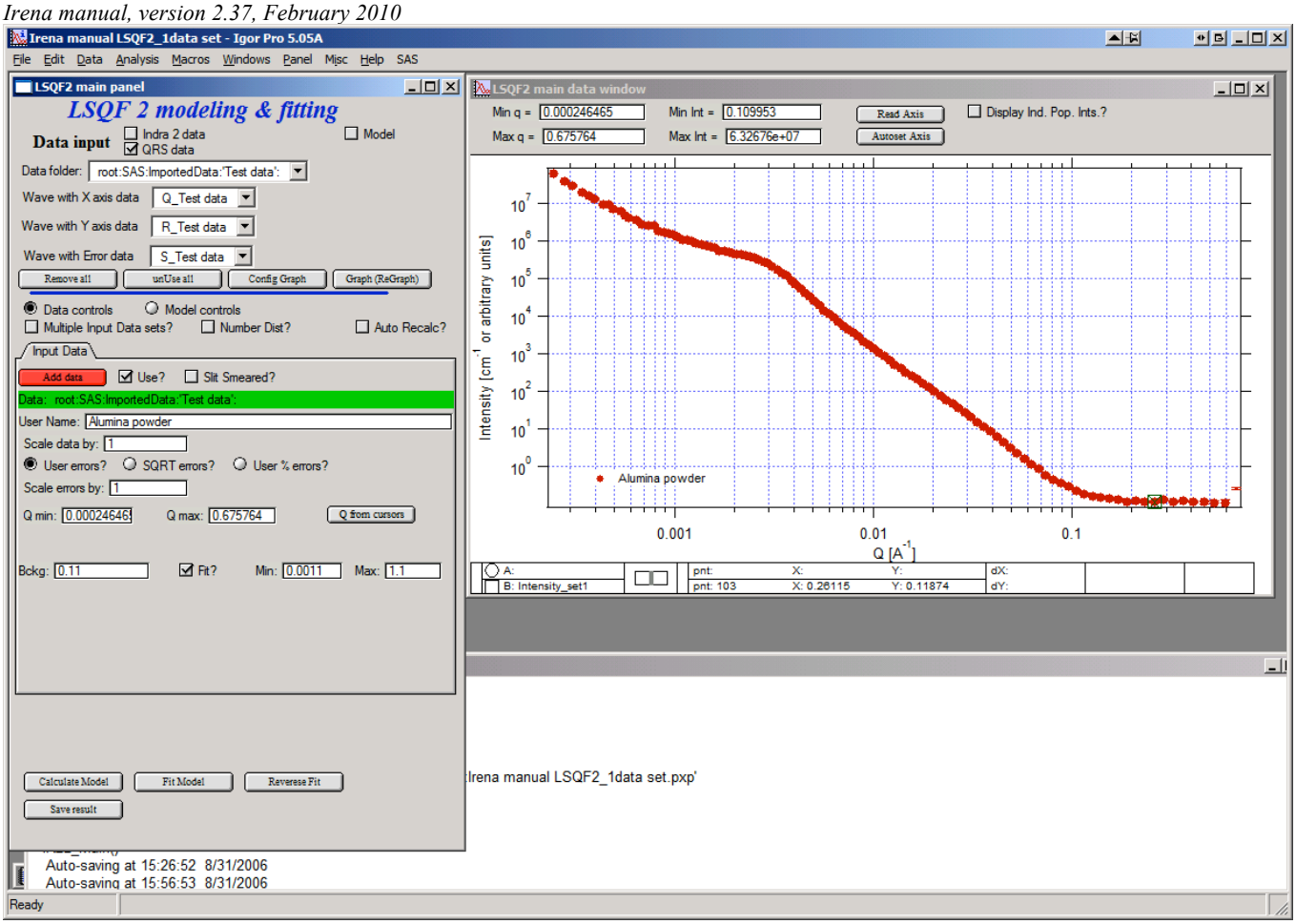
"Save result" Saves result into the folder. It feature is not fully finished yet. This is difficult to know what is expected in case of multiple data input…
"**Save in Waves"** Saves results into new folder in form meant for creating tables with results. In this case new folder (user is presented with dialog to create new name) is created and for each internal variable/string is created new wave. This creates large number of wavbes – most useless… But user then can create table of selected waves with important results – for example sample name, volume of pop1, mean diameter of pop1 etc.
"**Save in notebook"** Creates Igor Notebook (formatted) and pastes in this notebook summary of current state of the tool in more or less human readable form. This includes copy of the graphs and somehow reasonably formatted listing of parameters.

9.4 Fitting data with one input data set

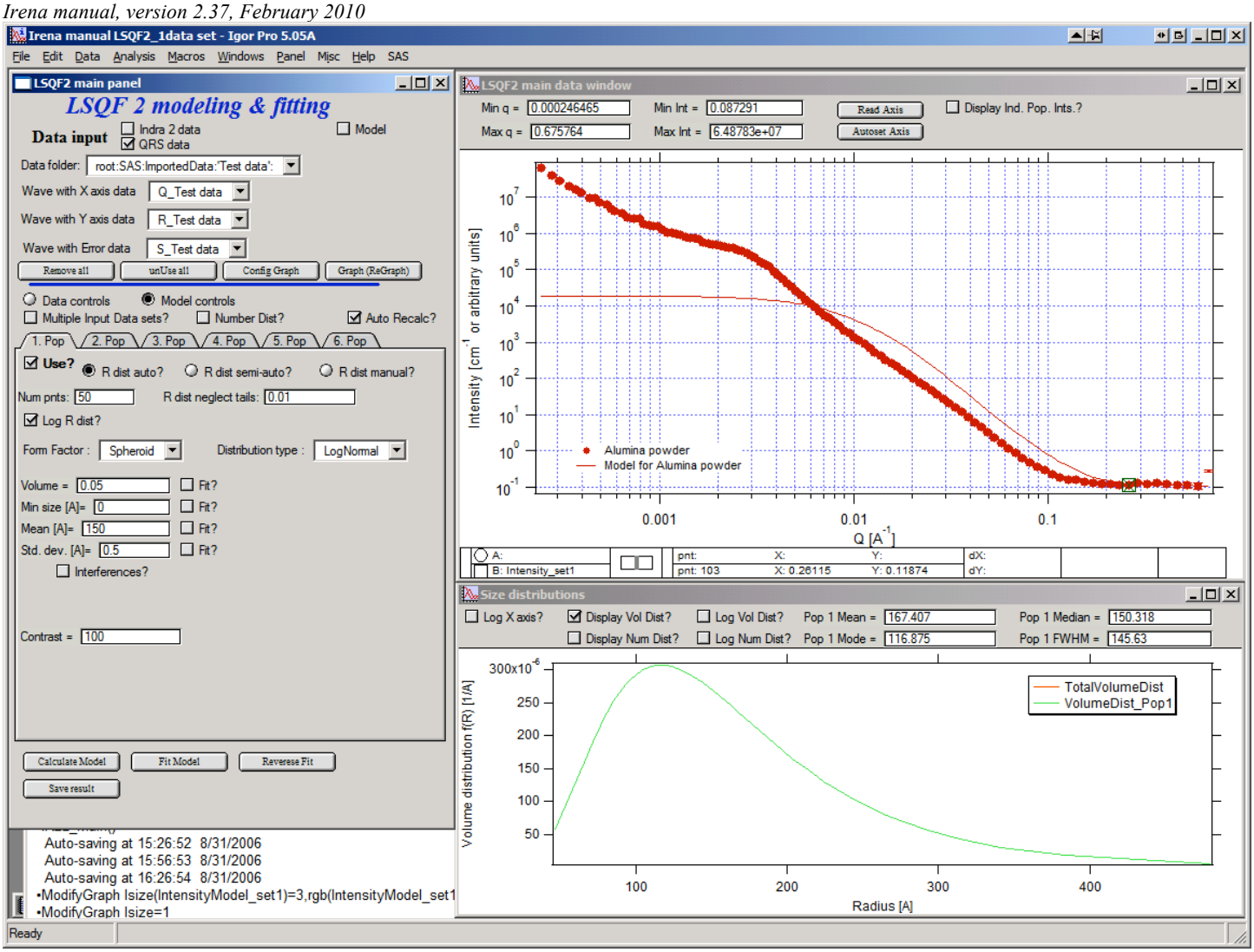
Select "data controls" radio button. Select data ('Test Data') and push red button "Add data"



Name the data “Alumina powder” in the “User name” field.
Let’s also select the background immediately here. Set cursor (square) to area of flat background (around point 100) and read value of intensity there from the reader below the graph. It should be around 0.12 or so. Type 0.11 into the “Bckg” field and check the “Fit?” checkbox. Note that the Min and max fields appeared and are set to 0.1 and 10x the value of our estimate. Uncheck the “Fit?” checkbox so the background is NOT fitted, when we run this next time…



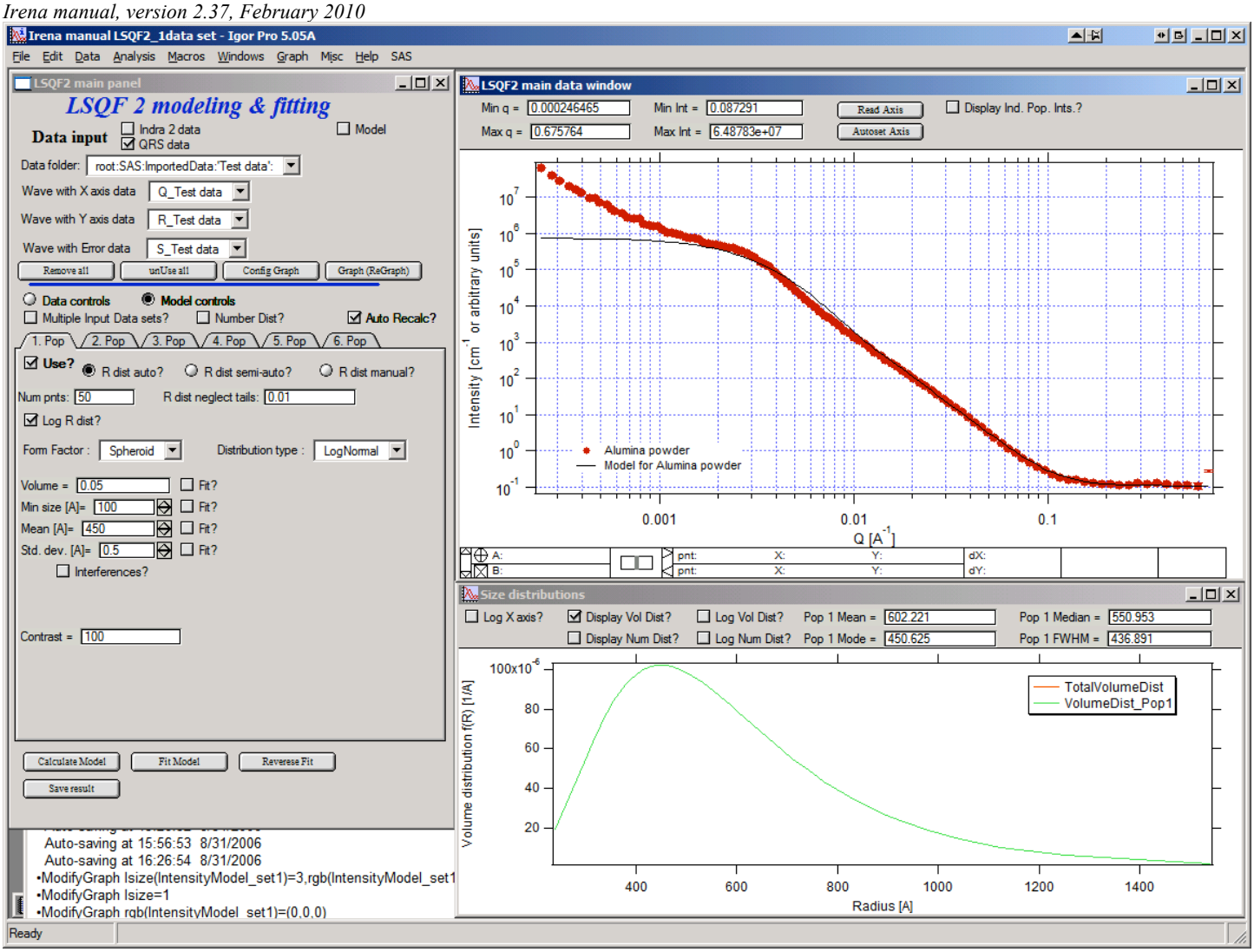
Now, let's go to "Model controls". Check the radio button "Model controls". Check the checkbox "Auto recalc". Make sure the "Interferences" checkbox is unchecked. Make sure that "Use?" checkbox for Pop 1 is checked and for all the others is unchecked.



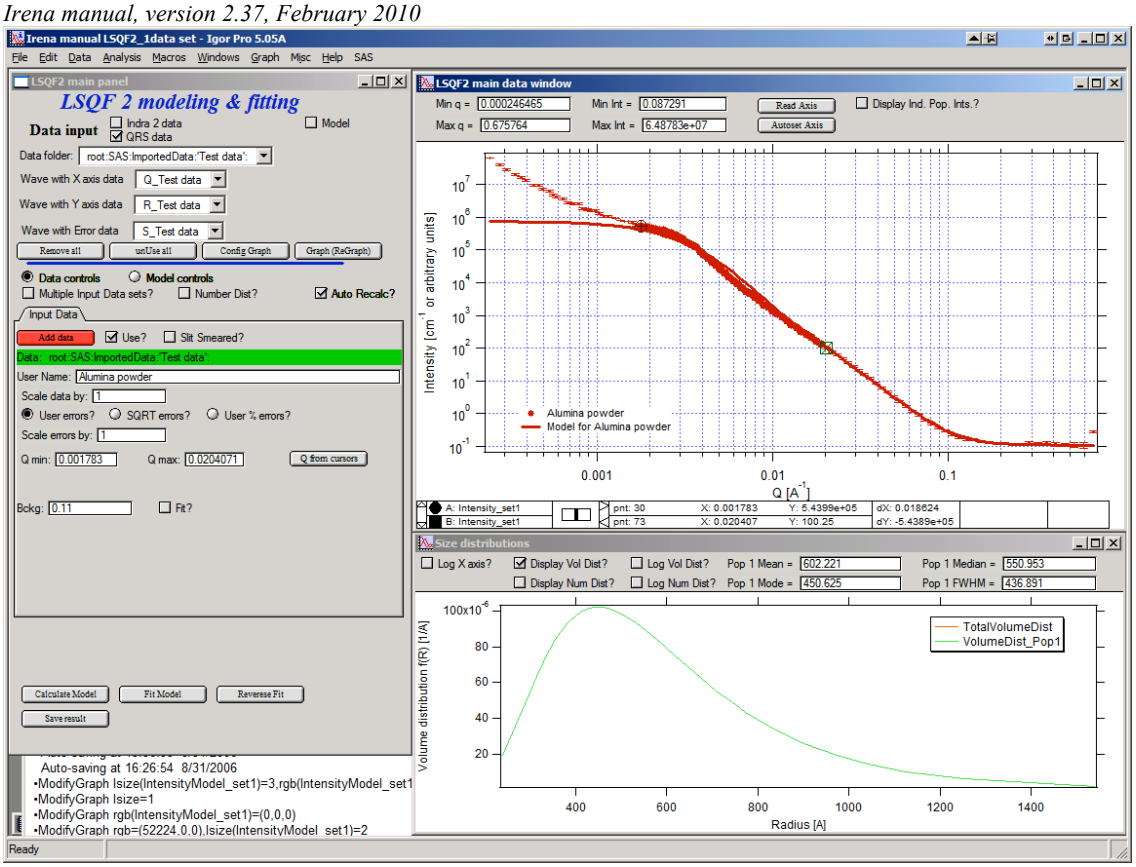
The model (default values) is going to be calculated.

Let's decide, that this population will be the larger stuff, dominating the data. This means the Guinier knee at around 0.003 A^{-1}. Also, since these data are not calibrated (powder sample), we can leave contrast to default value of 100. But if data would be meaningfully calibrated, correct contrast needs to be used here…

Let's change values little bit to get better estimate of parameters… Reasonable starting point is may be with Min size ~ 100, Mean ~ 450, and Std. dev ~ 0.5 :

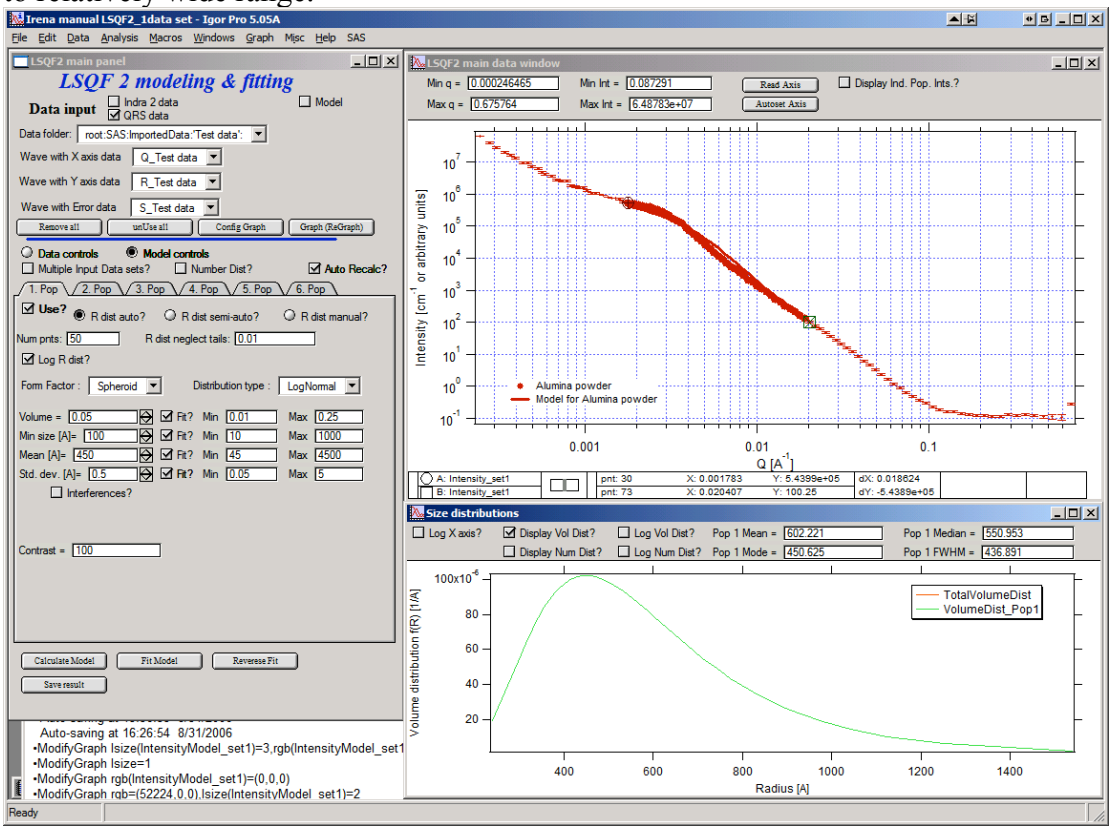


Now we need to select fitting range for this population… Check the "Data controls" Use cursors to select in the graph input data between point 30 and 73 and push button "Q from cursors". This will set the Q min and Q max values.

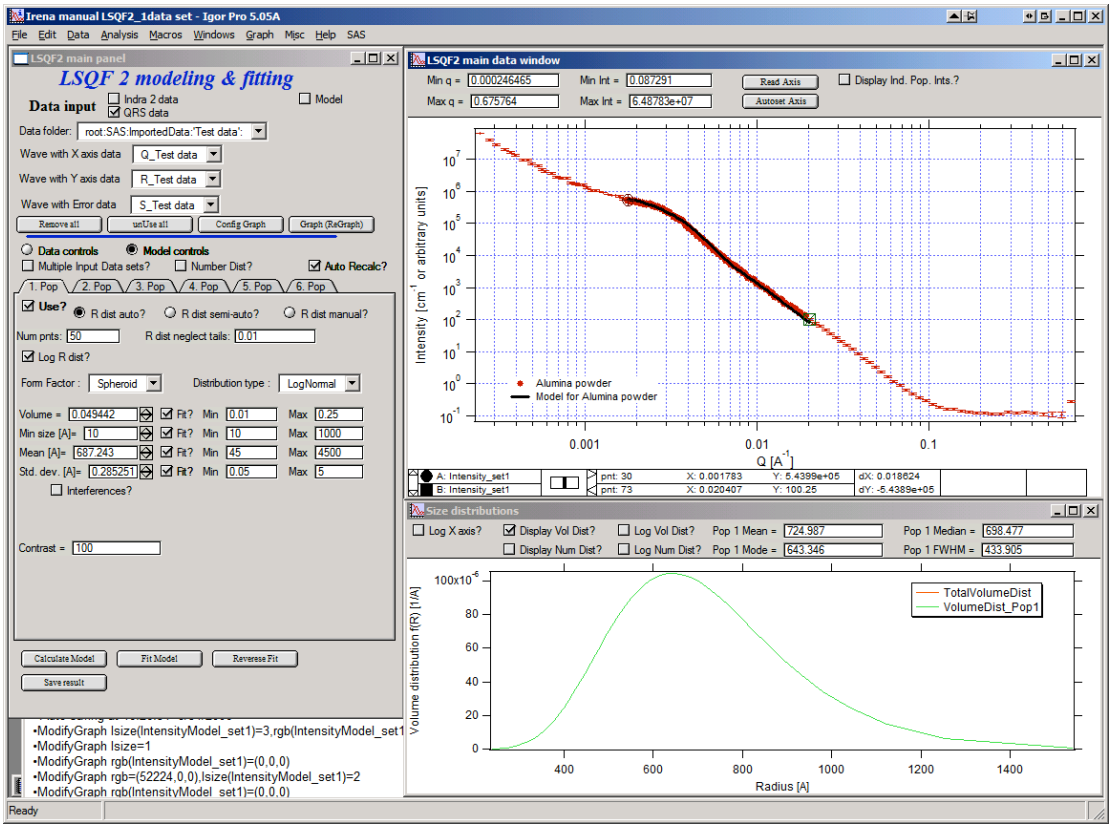


Make sure the background "Fit?" checkbox is unchecked here…

Now, let's select parameters to fit. Background is not appropriate for this subset of data. Select "Model controls" again. Check "Fit?" for Volume, Min size, Mean and Std Dev. Values for fitting limits should be set to relatively wide range.



Now push button "Fit model" at the bottom of the panel. The model should fit after few iterations…



Now we will add other population (smaller particles). Uncheck all "Fit?" checkboxes on this Pop tab.

Select Pop 2. tab. Check "Use?" checkbox here. To see whole q-range, go back to "Data controls" and change Qmin and Q max to smaller/larger values (0.0015 and 0.5). Now come back to "Data controls" and let's see, where the population 2 should be. Easiest achieved by unchecking "Use?" for population 1 and then the model in the graph is only for population 2. We want to use this population to describe data at around 0.05. So we need to move the mean to smaller sizes… This can be achieved by setting Mean to around 80 and reducing volume to about 0.01. Now check again "Use?" for population 1.

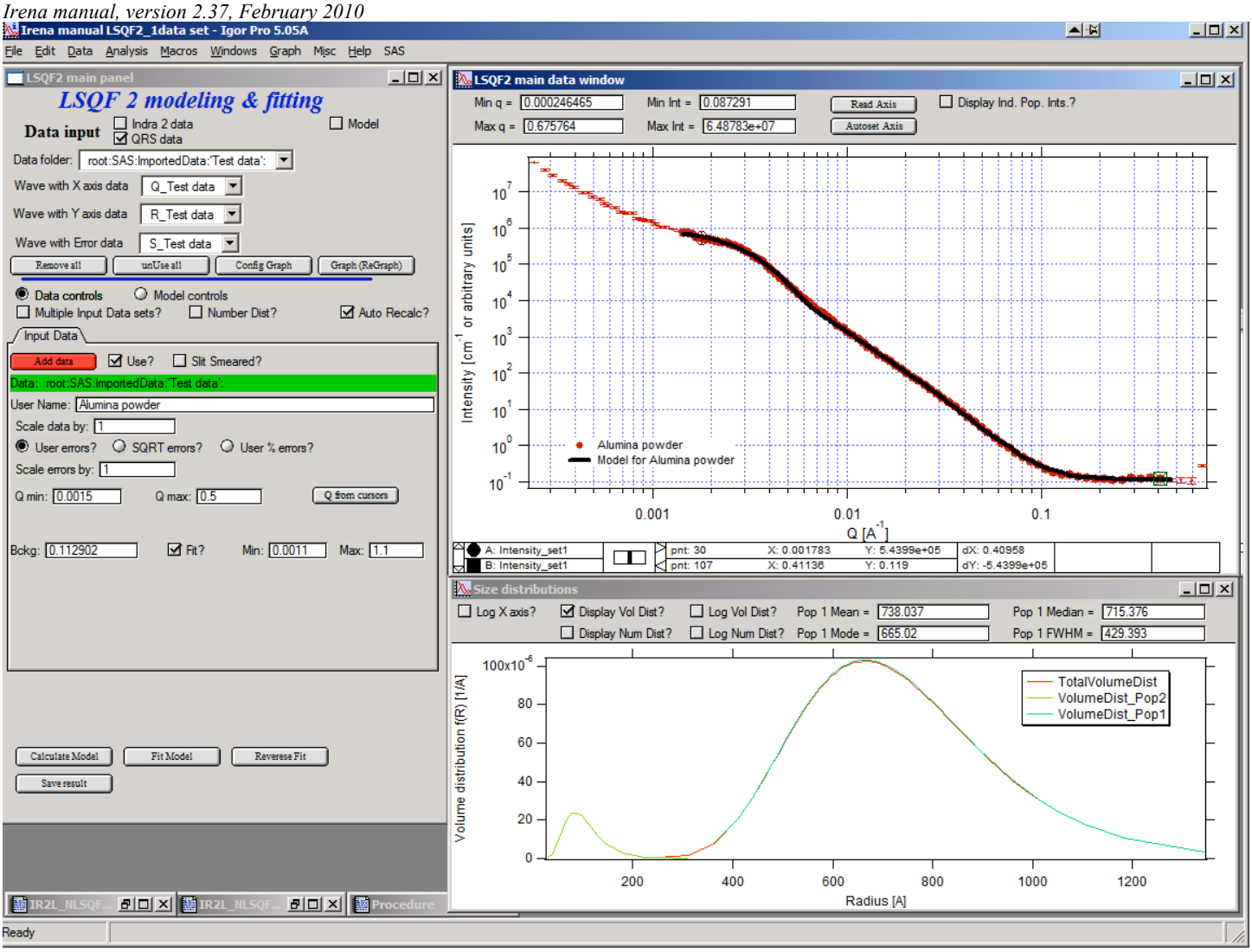
Check "Fit?" for Population 2 volume and Mean and fit the data by "FitModel" button.

Now we need to do final fitting of all meaningful parameters at once… There are now 3 places, where we need to select what will be fitted – but potentially could be even more… So let me review where the fitting parameters can be:

1. Data Controls – fitting of background
2. Model controls - Population tabs – fitting of distribution parameters (and volume, potentially interference parameters) – here we have two of these to check.
3. Model controls – Form factor panels - Potentially we could have for each population form factor parameters fitted, these need to be selected by reselecting again on each Population tab the form factor, which brings up (if appropriate) the appropriate panel..

Good luck finding all of the parameters… You need it.

Anyway, select background, Volume for each f the populations, Mean size, and Std deviation. Try to fit to the data from 0.0015A^{-1} to 0.5 A^{-1}… With little bit of luck (and a lot of calculations) you should get result similar to one below:



9.5 Fitting data with multiple input data set

Assumption of this chapter is, that you can already fit data with one data set (10.4). Only differences caused by adding other data sets are pointed out here.

Changes in Data controls

When "**Data controls**" AND "**Multiple Input Data Sets?**" are selected, up to 10 input data sets can be loaded at the same time in the tool. Each Data set has all of the controls as the first one, including separate background. Note, that if the background is to be fitted, checkbox needs to be selected on its tab.

If "**Different contrasts for data sets**" is selected, separate contrast needs to be input for every population and every population. This can be excessive number of contrasts. It is typically suitable for anomalous SAXS data evaluation.

Note, that the selection of number/volume distribution is used for all of the populations at the same type. You cannot mix number and volume distributions at the same time.

Note, that you can use one or more of the input data sets at the same time. If you unselect the "**Use?**" checkbox on any data tab, all parameters stay in the tab. Therefore you can mix-and-match data any time from any of the 10 populations.

Changes in Model controls

All controls stay the same. Contrast field will change reflecting selections: if "**Different contrast for data sets**" is not selected only one Contrast will appear, if it is selected, "**Contrast data X**" will appear, if Data X are set to be used. This appears on EVERY population tab. You need to go and check the contrasts for every population.

General comment

Please, remember, that with more data sets, this will be much slower. Setting up parameters for this complicated fitting space can be intimidating and very much complicated. You need to go through all of the used tabs in both Data controls and Model controls.

10. Size distribution (using maximum entropy, total-non negative least squares method and regularization)

10.1 Basic description of methods

Maximum entropy method by Pete Jemian

Maximum entropy (MaxEnt) and regularization (maximizes smoothness) are two separate methods for obtaining size distributions from small-angle scattering data. Yet, we describe them together here since they share many common components. Both are versions of a constrained optimization of parameters which solve the scattering equation.

$$I(Q) = |\Delta\rho|^2 \int_0^\infty |F(Q,r)|^2 V^2(r) NP(r) dr \quad (9.1)$$

The difference in these two methods is in the applied constraint and it is this constraint which most heavily influences the differences between the two methods in the form of the result.

The maximum entropy method was developed by Jennifer Potton et al., and supplied in the code package `MAXE.FOR`. Pete Jemian (jemian@anl.gov) has had his hands all over this code and in a few places, made some rather significant additions, resulting in the code package `sizes.c`. Most significant is the addition of the regularization method which is likely to succeed an finding a solution in many cases when the MaxEnt method fails to converge upon a solution. Please contact him with any questions regarding the implementation of these methods. (Point of fact, *both* are actually regularization methods.)

J.A. Potton, G.J. Daniell, and B.D. Rainford; ***Inst Phys Conf Ser* #81, Chap. 3** (1986) 81-86
--- ***J Appl Cryst* 21** (1988) 663-668
--- ***J Appl Cryst* 21** (1988) 891-897.
J. Skilling and R.K. Bryan; ***Mon Not R Astr Soc* 211** (1984) 111-124.
Ian D. Culverwell and G.P. Clarke; ***Inst Phys Conf Ser* #81, Chap. 3** (1986) 87-96.

Literature citation for Maximum Entropy code in Irena macros bvy Pete Jemian

Pete R. Jemian, Julia R. Weertman, Gabrielle G. Long, and Richard D. Spal; Characterization of 9Cr-1MoVNb Steel by Anomalous Small-Angle X-ray Scattering, **Acta Metall Mater 39** (1991) 2477-2487.

Here *NΠ(r)* is described as a histogram size distribution where a fixed number of bins are defined over a given range of diameter with either constant diameter bins or constant proportional diameter bins. Solution of the histogram size distribution to the scattering equation 9.1 above is obtained by fitting the scattering calculated from trial distributions to the measured data and then revising the amplitudes of the trial histogram distribution based upon the applied constraints. The trial histogram size distribution is not forced to adhere to a particular functional form, such as Gaussian or log-normal. However, in the current formulation, all sizes of the scatterer are expected to have the same scattering contrast and morphology (shape, degree of interaction, aspect ratio, orientation, etc.).

In both MaxEnt and regularization methods, the measured data must be represented by the calculated data so that the goodness of fit criteria (sum of squared standardized residuals) is close to the number of measured data points used in the analysis, subject to an additional constraint. This imposes a high standard for the reported errors on the scattering intensity. The reported errors are expected to be estimates which are comparable to one standard deviation of the true intensity and that the difference between the measured intensity and the true intensity is within one standard deviation of 67% of the time and randomly distributed such that a summation over these differences has zero mean and unit RMS. If these conditions are not met, it is likely that artifacts in the derived size distribution will result. Often it is necessary to scale the reported errors by a factor to achieve converge of the MaxEnt method.

As a point of fact, both MaxEnt and regularization are regularized methods of solution to the scattering equation (9.1). They both seek solutions of the functional, Ξ,

(9.2) $\Xi = \chi^2 - \alpha S$

where χ^2 describes the goodness of fit, S is the applied constraint, and α is a Lagrange mutiplier used to ensure that the solution fits the measured data to some extent.

For MaxEnt, the additional contraint is that the configurational entropy of the size distribution must be maximized. Rather than be bothered by what this means when compared with the thermodynamic entropy, you are asked to consider that this constraint enforces the principle that all histograms in the size distribution must have a positive amplitude. To make the calculation of the entropy, an additional reference level must be defined. Typically, this reference level (a.k.a., Sky Background, starting guess, *a priori* information) is about 0.01 of the maximum level of the final size distribution. One does not need to fine-tune this parameter and should never be concerned with adjustments less than one order of magnitude. Too high and this parameter will cause the solution to have upward tails at both low and high ends of the distribution. Too low and additional scatter will appear in the distribution. The MaxEnt constraint imposes no correlation on the amplitudes of adjacent bins in the calculated histogram size distribution.

Regularization method by Pete Jemian

The regularization method implemented here maximizes the smoothness of the calculated histogram size distribution by minimizing the sum of the squared curvature deviations. The particular mathematics used here do not prevent the use of negative values for the amplitudes of the histogram size distribution and this is a noted behavior which must be considered to avoid. Often, it is possible to avoid the negative bins in the size distribution by adjusting the fitting range, the bins in the histogram size distribution, or the background.

Total non-negative least square method

This is implementation of the "Interior point method for totally nonnegative least square method". I have found reference and method description for this method on line: Michael Merrit and Yin Zhang, Technical report TR04-08, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, 77005, USA. This publication was from May 2004, I have found it on the web posted in December 2004, http://www.caam.rice.edu/caam/trs/2004/TR04-08.pdf
Basically, this is very interesting method, in which one starts with reliably positive solution, calculates gradients using least square method to better solution and makes step towards this solution. The size of the step is limited in such manner, that the solution (histogram bin content) cannot be made negative. If the step would make it negative, the size of the step is limited in such manner, that the non-negativity is guaranteed.
The problem of this method is, that there does not seem to be any simple way of incorporating errors in the calculation. Generic method which was suggested to me resulted in instability of the code. So, contrary to MaxEnt method (which inherently uses errors), in this method the errors are used only to identify sufficiently good solution.
Also this method seems to have major problem with the poor conditioning of the SAS problem – natural log-q and log-I behavior of the SAS data. Therefore, it basically requires, that fitting is done in different "weighing" of the data – for example $I*Q^4$ vs Q etc…

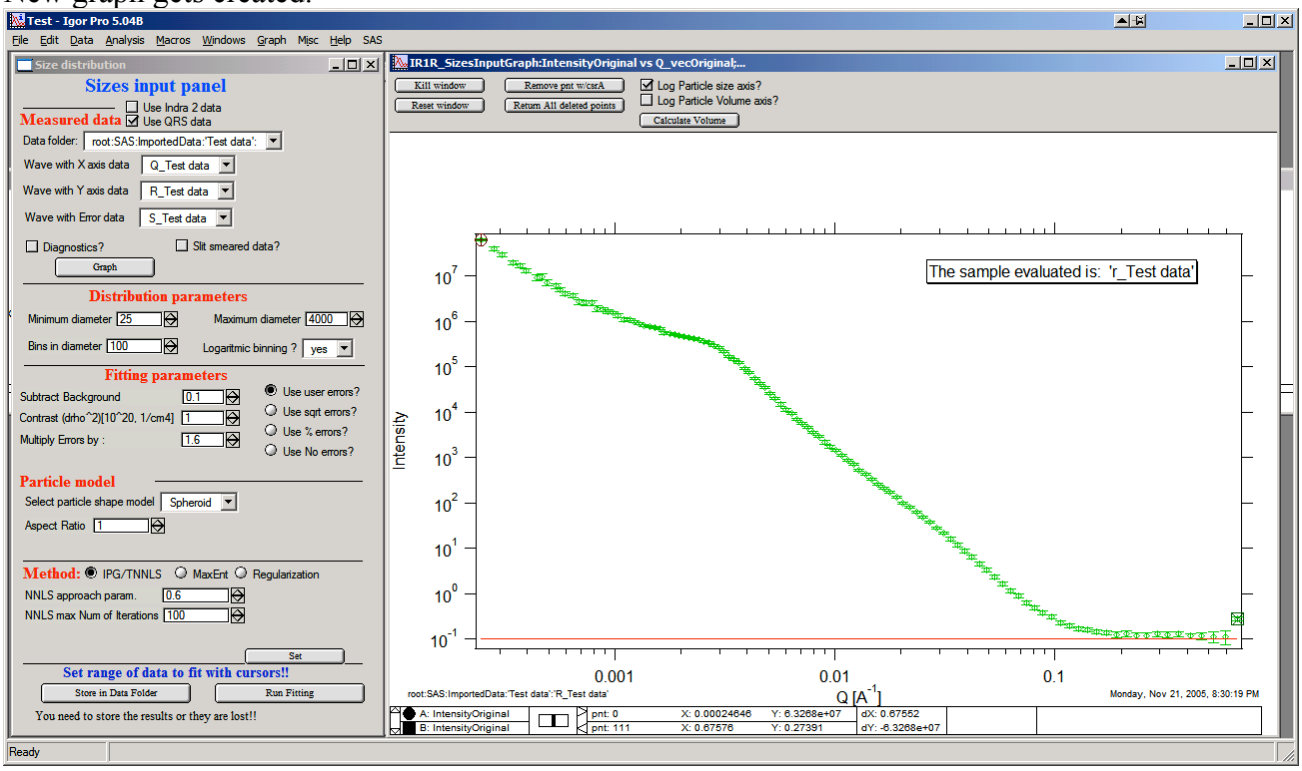
10.2 Using the tool

This program uses one complex interface – a complex graph and panel for data input and manipulation. To start, select "Size distribution" from "SAS" menu…

On the panel, which gets created, starting from top:

1 select the "Use QRS checkbox" (assuming you are using QRS named data as explained above).
2 Select data folder with data (see figure below)
3 Select wave with Q vector, other should be selected automatically (if not select right waves). Note, that it is now not necessary to input error wave. See below…
4 "Graph"

New graph gets created.



Leave the "Slit smeared data" set to no and "Slit length" set to NaN. If using the Indra data structure (UNICAT USAXS data reduction), these fields are preselected in the proper form and should not have to be changed. If the data are from different instrument (as here) and are slit smeared the macro can be still used. Providing user selects correctly slit smeared data and inputs slit length in units of Q. I expect this case to be highly unlikely…

Next we need to setup some parameters.

Distribution parameters:

Minimum diameter & Maximum diameter – both are in A. These are limits of fitted distribution. Set minimum to 25 and maximum to 10000

Bins in diameter – into how many bins you want to divide the range of diameters. 100 is a good number – more points may be really slow on slower computers.

Logaritmic binning – if yes, the bins are binned logarithmically – i.e., the bins at small sizes are smaller and at large sizes are larger, giving save width bins when plotted on axis logarithmically. This is very useful setting for the wide ranges of sizes measured using USAXS instrument. If no is selected here, the bins are all same width. Leave in yes for now…

Fitting parameters:

Background this is flat background to be subtracted from data. The red line in the graph shows current value. Set correctly for this case to 0.1 or so

Contrast (delta rho squared) – if this is properly inserted, the data are calibrated… Leave to 1 since the contrast is not known.

Error handling:

There are four ways to handle now errors in this tool. The method is selected by four checkboxes lined vertically next to the "Background and Contrast" fields…

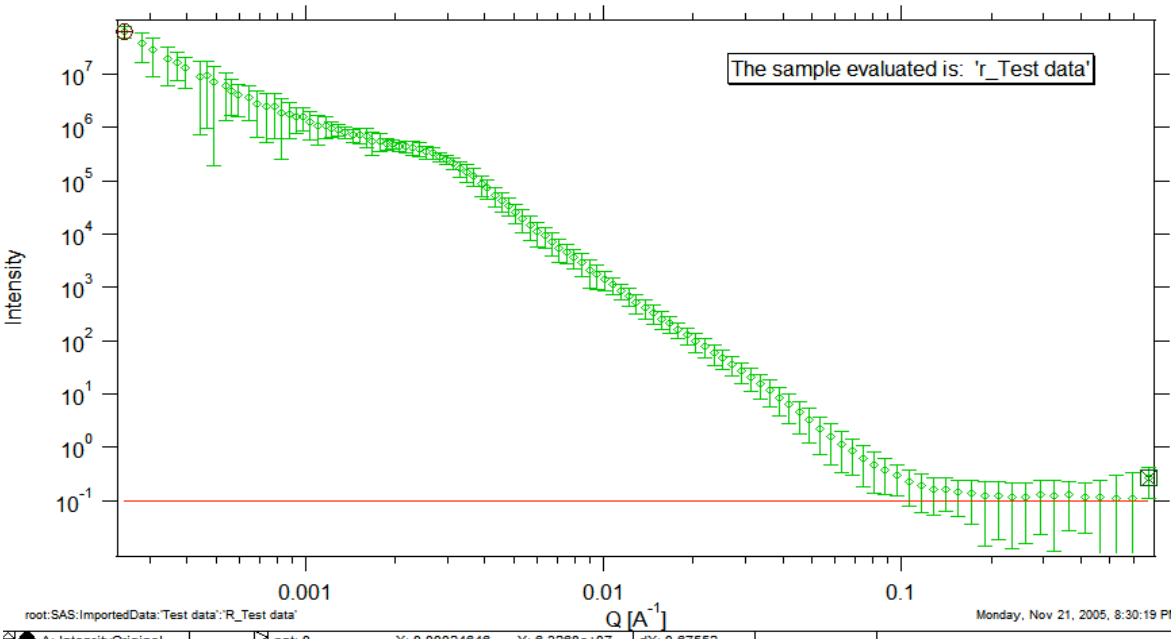
1. "Use user errors" use erros input as wave. In this case the field: "Multiply errors by"is available and errors can be scaled as needed. Start with high multiplier and reduce as necessary to reach solution, which is both close to the data but not too noisy.
2. "Use sqrt errors" – will create errors equal to square root of intensity (standard Poission error estimate). You can multiply these errors by error multiplier. Errors are smoothed.
3. "Use % errors" – will create errors equal to n% of intensity. Field where to input the n appears. Errors are smoothed.
4. "Use No errors" – use no errors – the weight of all points is the same. This is unlikely to be correct, but this case allows to use fitting in "scaled" space – Intensity * Q^m vs Q, where m = 0 to 4. This helps to mathematically better condition problem (similarly to using errors) and can yield sometimes good solution.

NOTE : at this time you cannot use this method (no errors) with MaxEnt or Regularization.

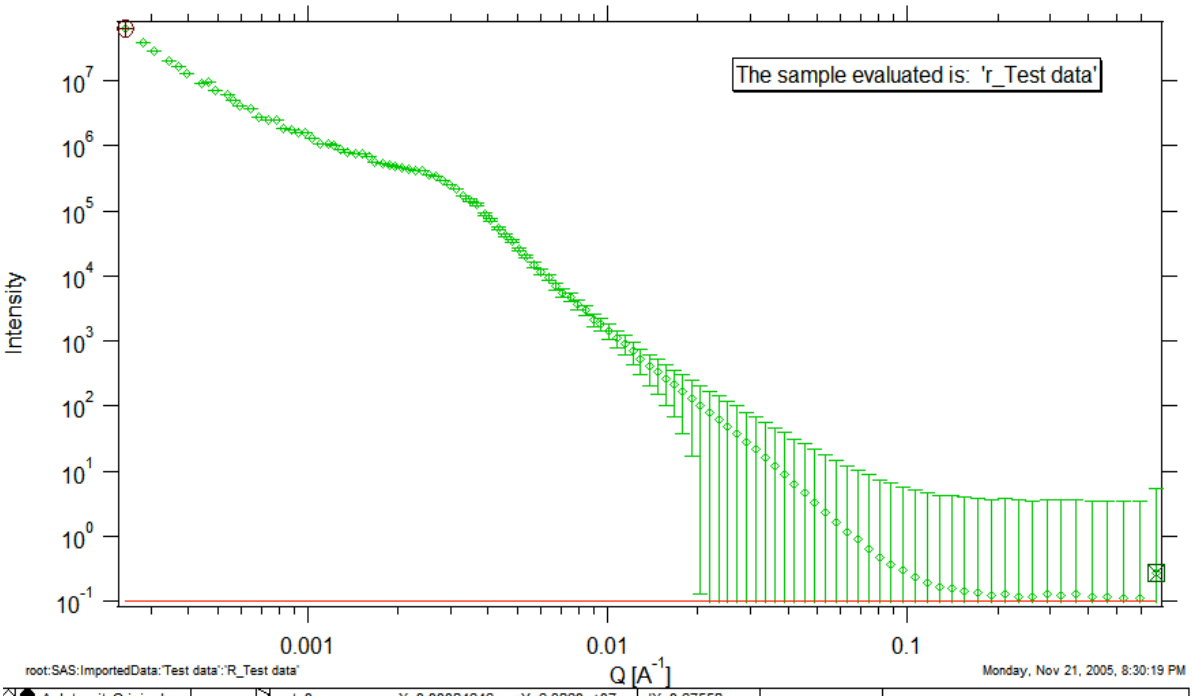
Comments:

MaxEnt works best with user errors or % errors.
IPG/TNNLS seems to work best with no errors and m = 2 -4. Reason is unclear.

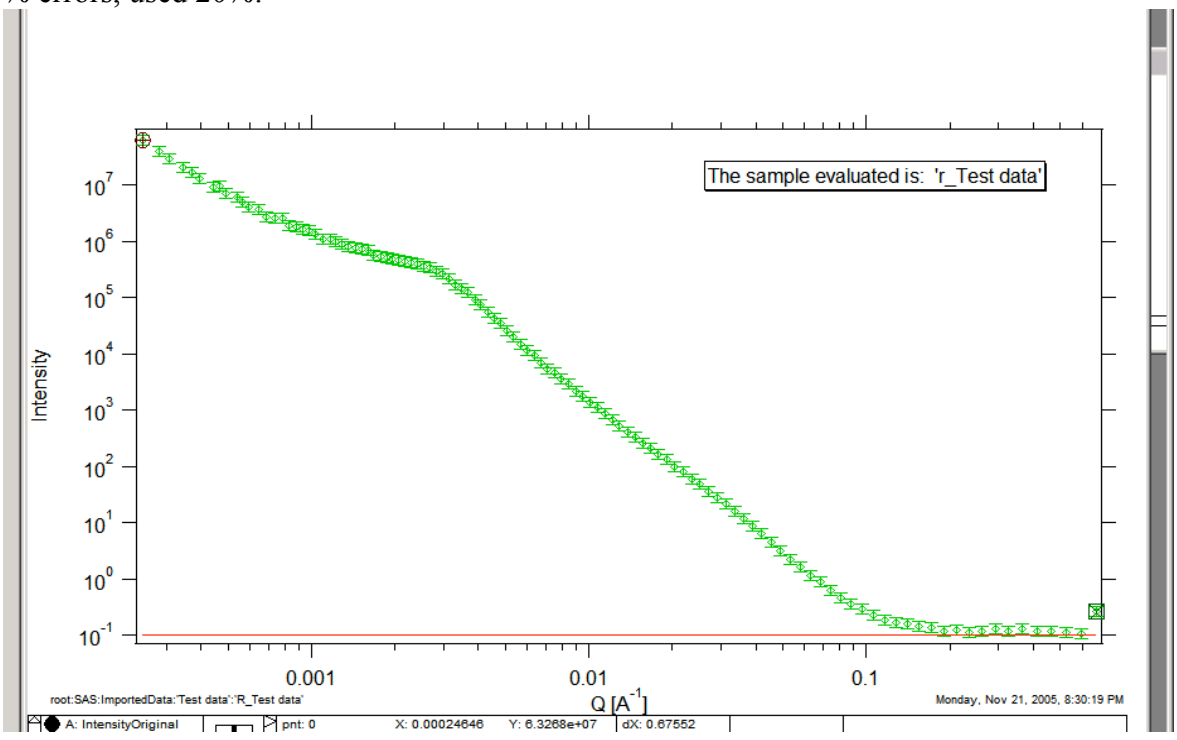
The errors displayed in the graph will change as different methods are selected:
User errors, multiplied by 10:



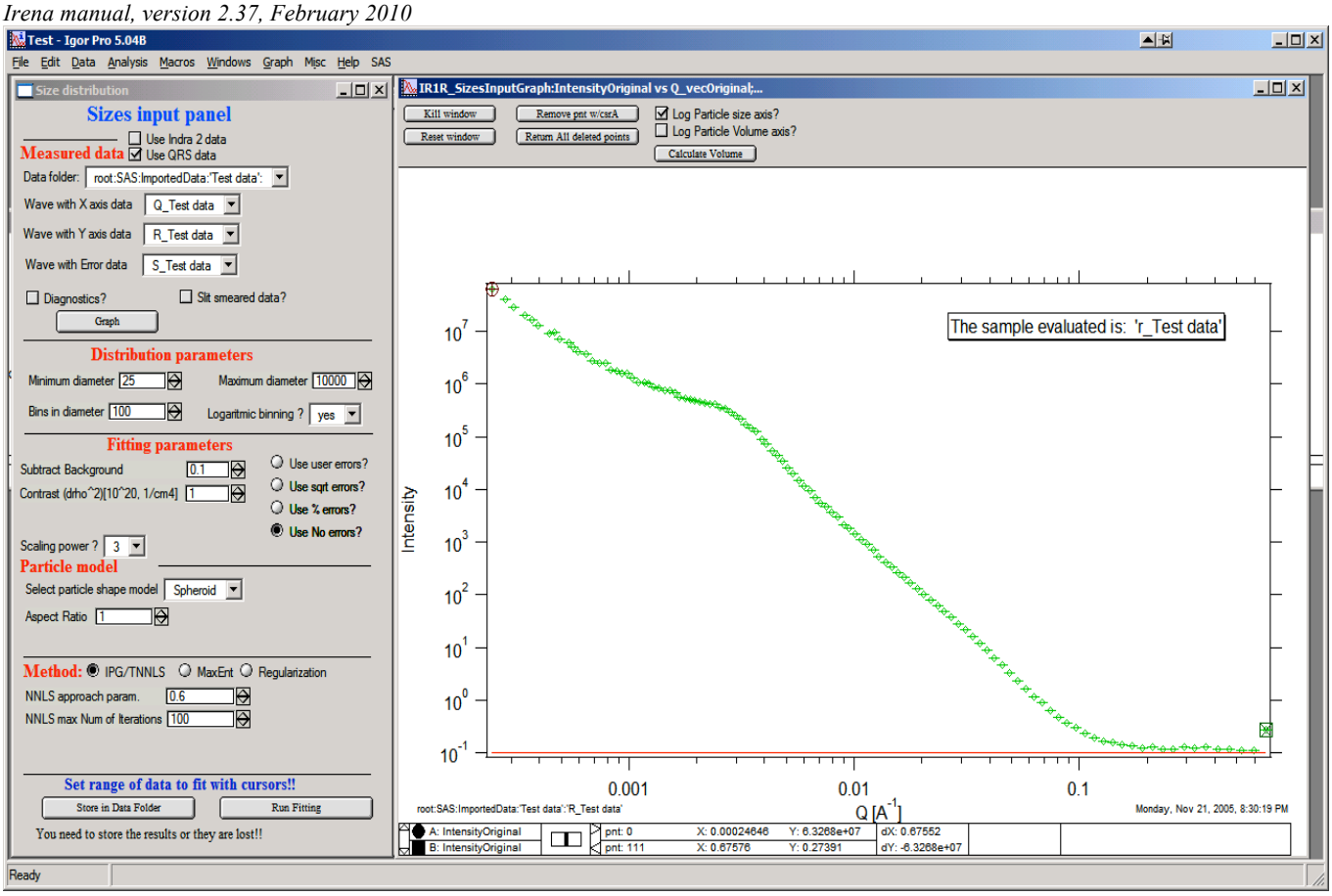
SQRT errors, multiplied by 10:



% errors, used 20%:



No errors, selected to use I*Q^3 vs Q "space" for fitting:



Particle shape:

Particle shape model – the tool uses the same selection of form factors as Least square fitting. If you feel you really need another shape, I can put it in. Same comments apply WRT speed as mentioned in Least square fitting – "integrated spheroid" is using the most complex way to avoid possible artifacts, but is very slow. Spheroid AR 1 is fastest, others depend on complexity of math and integration. The code has been internally optimized to run as fast as possible.

Aspect ratio – anything, 1 is for sphere.

Method:

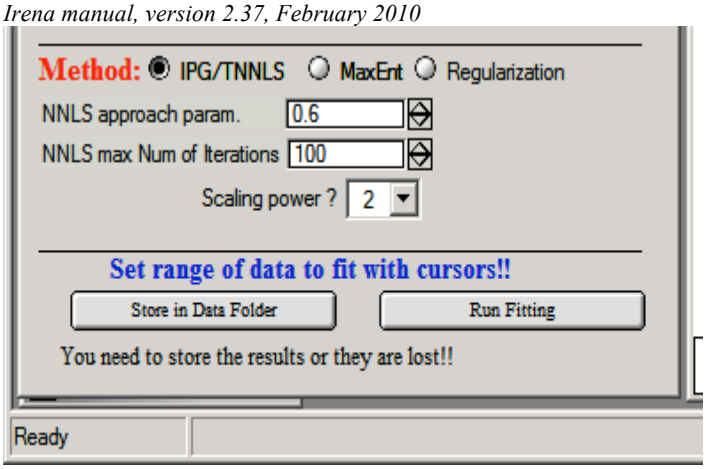
The default method is Maximum Entropy.

Size precision parameter is internal number which should not be changed too much. Most users should be happy with default. Smaller the number, more precisely MaxEnt needs to match the chi squared…

MaxEnt max number of iterations – unlike Regularization, which has limit on number of iterations, MaxEnt can go infinitely. Therefore maximum number of iterations need to be enforced.

MaxEnt Sky Background. While this is relatively complicated number internally, note the suggestion next to it. Suggested value is 0.01 of maximum of the resulting volume distribution. The suggested value will be either green or red, depending if the value in the box is reasonable. Accept the suggestion and you will be happy.

IPG/TNNLS



Approach parameter is the step size (from maximum) which will be made in each step towards calculated ideal solution. Basically convergence speed, but too high number will cause some overshooting and oscillations. For most practical purposes seems to work fine around 0.5-0.6.
NNLS max number of iterations – limits number of iterations. Change as needed.
Scaling power – this is how Intensity will be scaled to improve the conditioning of the problem.

Regularization
Has no additional controls.

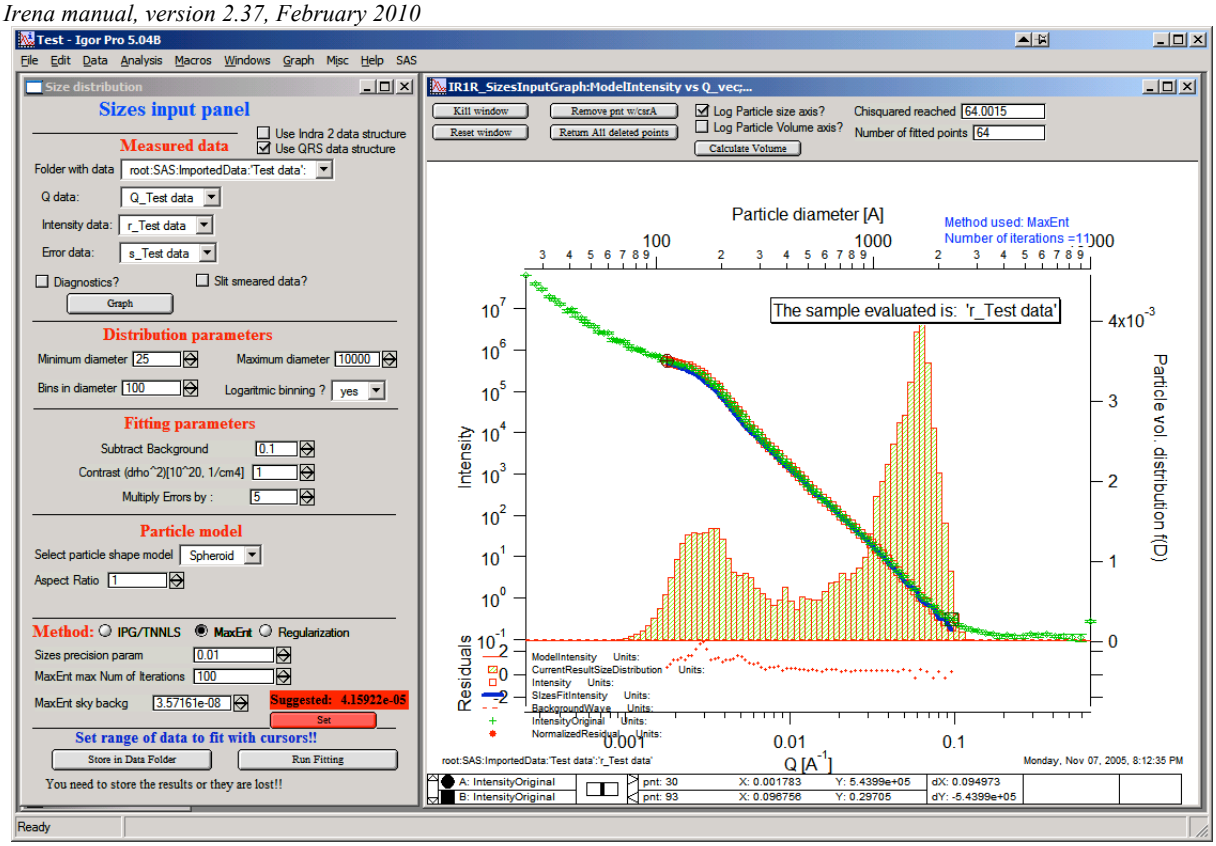
Buttons part

"Run fitting" runs the above selected method.

SAVE THE RESULTS button – if you do not push this, the data are not copied back into the sample folder and are overwritten with new data.

Getting fit.

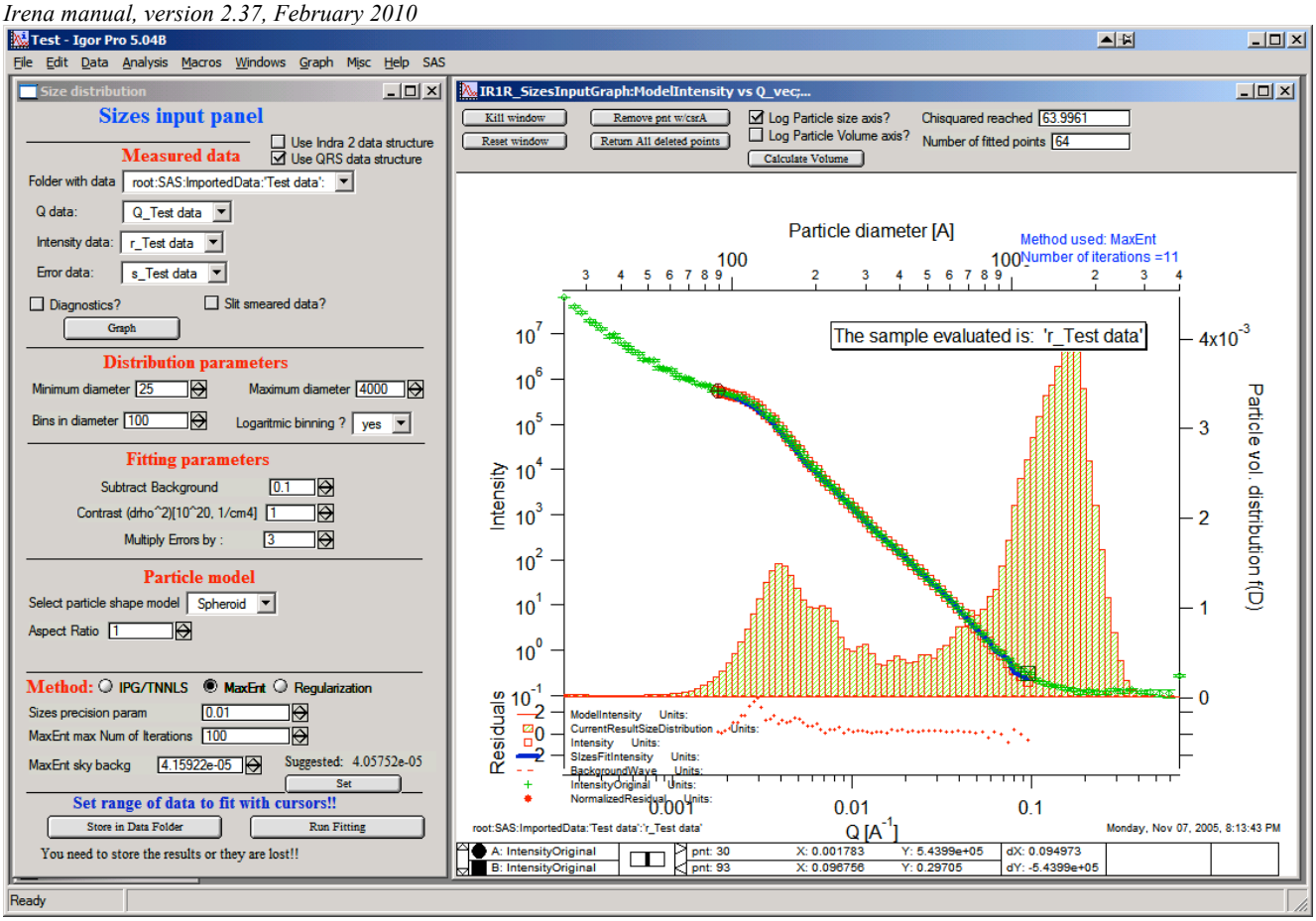
First select range of data using the cursors. Set rounded cursor on point about 30 and squared on point 89 or so. Note, that you can vary the range of fitted data between the fits.
Push button "Run internal MaxEnt". Solution should be found as in the figure below…
If the parameters are too restrictive you may get error message, that solution was not found. In such case check minimum and maximum diameter settings, check the error multiplication factor etc. Generally I suggest starting with higher range of radii than needed and higher error multiplication factor. Then reduce as needed. Also check the shape.



This is rough fit for the data in the graph – and for purpose of description of this graph now.

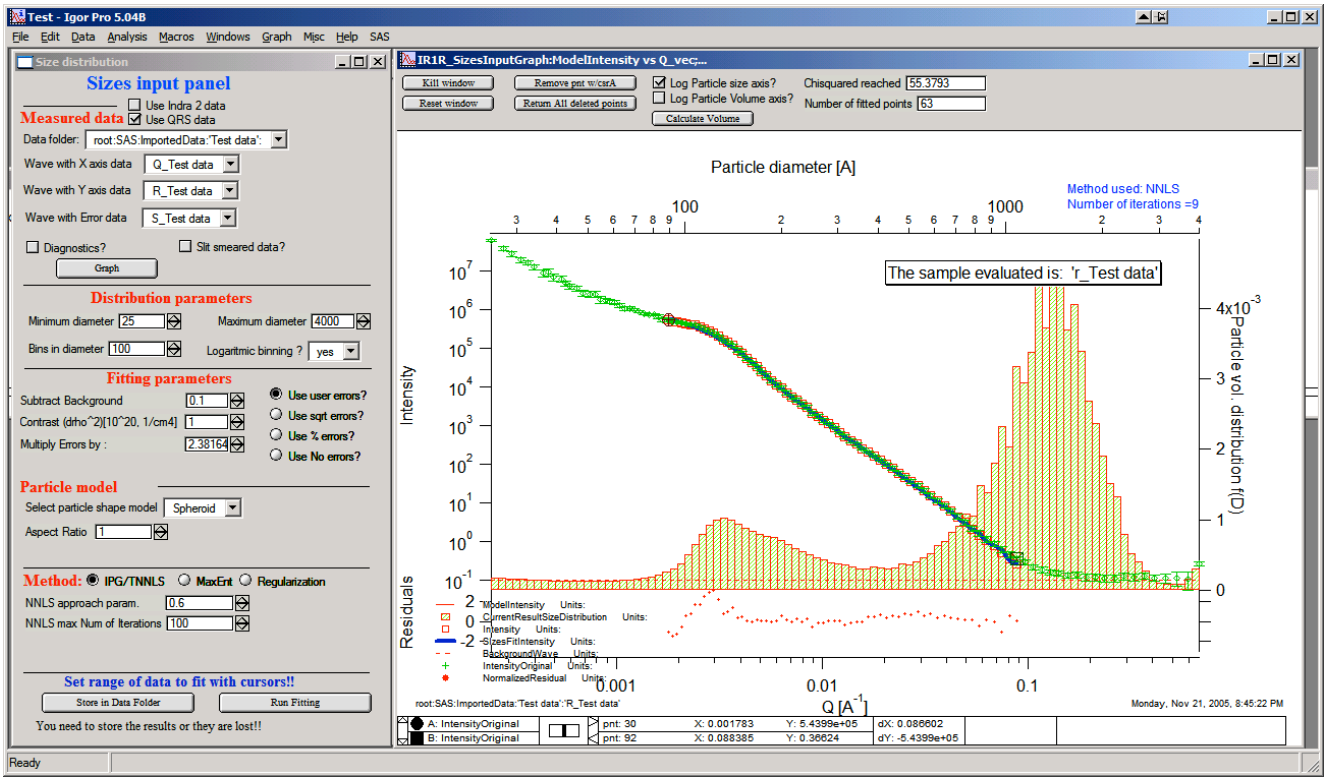
Now let's get to explanations:
The green points are the original data points.
The red points (top part of graph) are points selected for fitting (without background)
The blue line (very difficult to see) is the fit obtained by the fitting routine
The bar graph is the particle volume distribution (use top and right axis)
In the low graph
The red dots are normalized residuals. Ideally these should be random within +1 and –1, this structure suggests some misfits in some areas.

To get better results one now needs to play with the parameters. I suggest reducing maximum diameter to about 4000A, reducing multiply errors by to 3, fixing the MaxENt sky background and the running the same routine again. Following is the result:

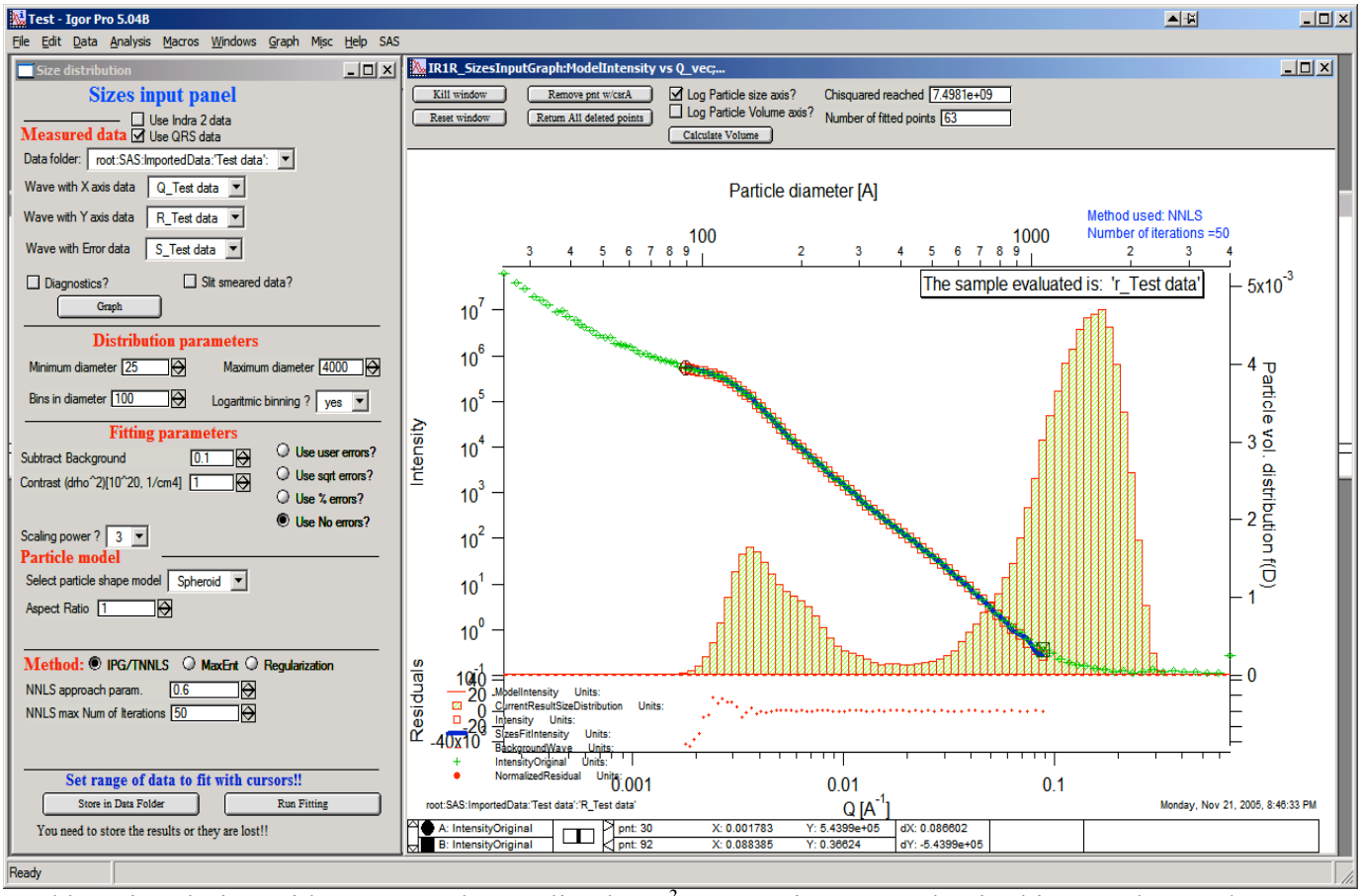


This shows, that we have bimodal distribution of scatterers. By the way, these data are from mixture of two polishing powders.

And now the IPG/TNNLS method:



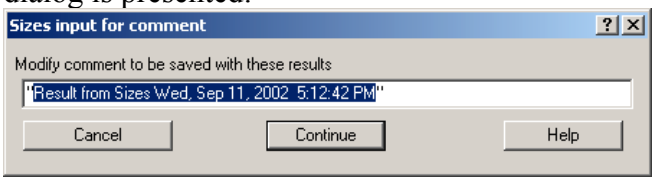
This is solution with user errors. Note, that the solution is basically very similar to MaxEnt.



And here is solution with no errors, but scaling by Q^3. Less noisy. Note, that in this case the IPG/TNNLS method is stopped by the Maximum number of iterations. Less number of iterations, less noisy solution – but may not be close to measured data…

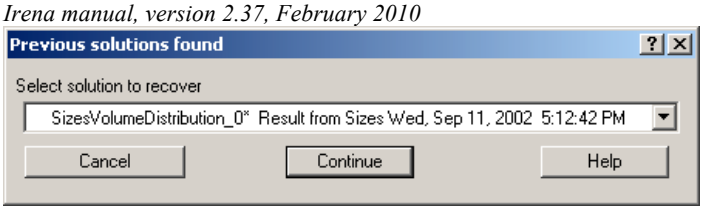
NOTE : at this time you cannot use this method (no errors) with MaxEnt or Regularization.

Saving the data copies waves with results into folder where the measured data originated. Also, it is possible to have various generations of data saved. In order to give user chance to find what each saved result is, following dialog is presented:



Here user can write ANYTHING, as long as it is bracketed by the QUOTES. The QUOTES are VERY important.

If user tries to start Size distribution macros in folder, where saved solution to this method exists, he/she is presented with dialog, which allows one to recover most of the parameters used for that solution.

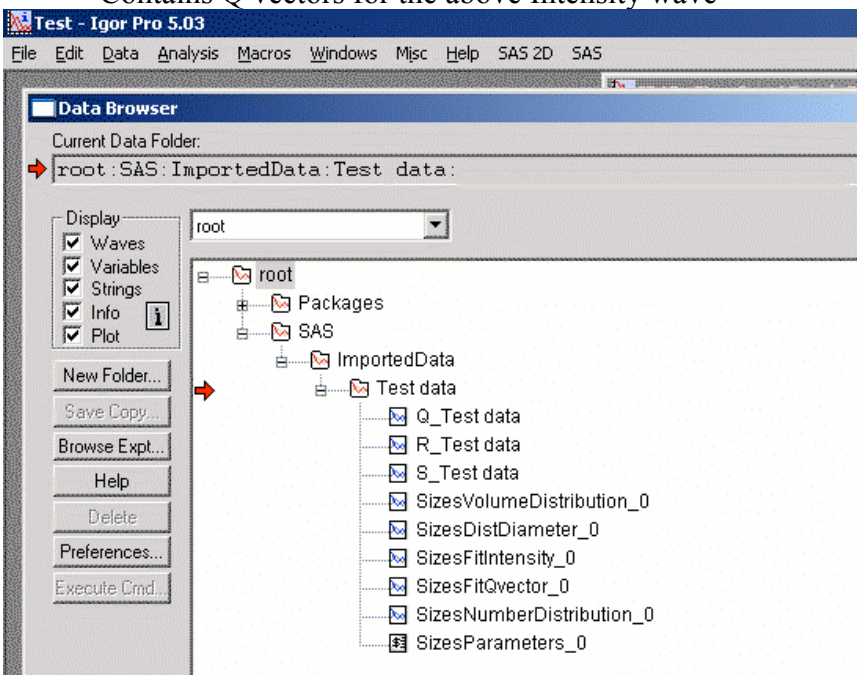


Therefore it is possible to start from where he/she left off. Also it is possible to start fresh - just hit cancel in this dialog - when parameters are left in the state they are left in after last fitting (or in default if this macro was not yet run in this experiment.

Resulting waves:

Following waves are created in the folder with data, when saved from this macro (_0, _1, _2, etc are different generations of solutions saved by user):

SizesNumberDistribution_0
Contains number distribution data
SizesVolumeDistribution_0
Contains volume distribution data
SizesDistDiameter_0
Contains Diameters for the other waves which need it
SizesFitIntensity_0
Contains Intensity of the model
SizesFitQvector_0
Contains Q vectors for the above Intensity wave



Comment, each of these waves contains WaveNote (see below at the bottom of the figure), which contains most of the details about how the particular results were obtained:

These are the parameters:
SizesDataFrom=root:'Test data':
SizesIntensity=Intensity

SizesQvector=Qvector
SizesError=Error
RegNumPoints=40
RegRmin=12.5
RegRmax=2000
RegErrorsMultiplier=3
RegLogRBinning=yes
RegParticleShape=Spheroid
RegBackground=0.12
RegAspectRatio=1
RegScatteringContrast=1
RegSlitSmearedData=No
StartFitQvalue=0.001783
EndFitQvalue=0.068163
RegIterations=12
RegChiSquared=60.45
RegFinalAparam=1.8853e+07
UsersComment=Result from Sizes Wed, Sep 11, 2002 5:12:42 PM
Wname=SizesDistributionVolumeFD_0
Most of these parameters should have self explanatory names. This is where user can figure out what happened. Further some parameters are also saved in the string with name "SizesParameters_0" such as MeanSizeOfDistribution.

11. Pair distance distribution function (PDDF, p(r))

11.1 Model description

This tool calculates Pair distance distribution function as generally defined in the small-angle scattering theory (see any basic SAS book, like Glatter/Kratky 1982, page 27, formula 29):

$$I(Q) = (\Delta\rho)^2 V \int_0^D 4\pi r^2 dr \gamma_0(r) \frac{\sin(Qr)}{Qr} = (\Delta\rho)^2 V \sum_0^D 4\pi r^2 \Delta r \gamma_0(r) \frac{\sin(Qr)}{Qr}$$

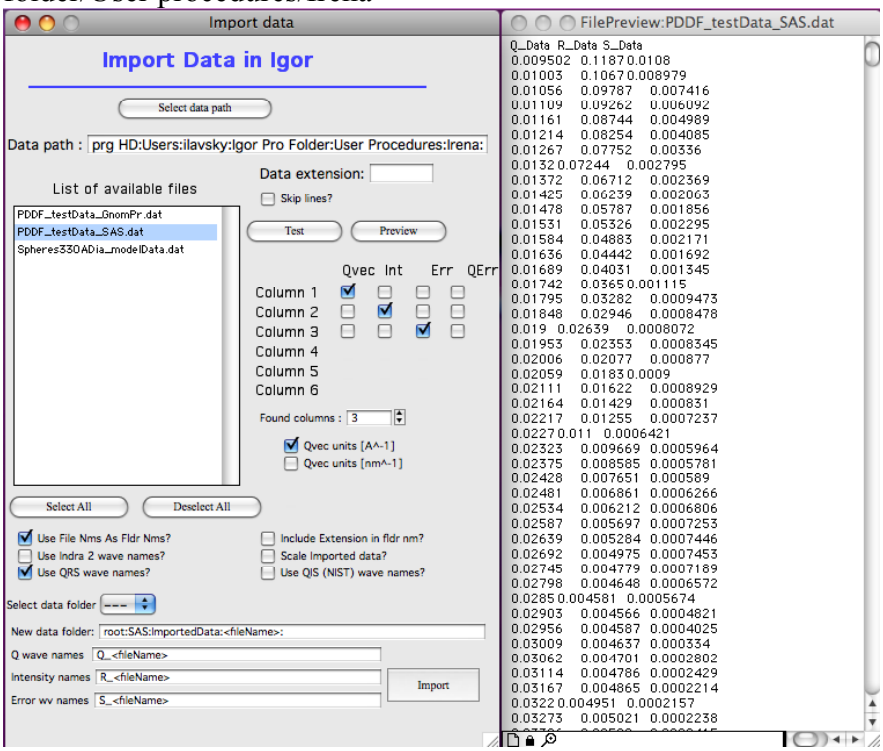
where the $(\Delta\rho)^2 V$ is contrast/volume of the scatterers (simply scaling) factor. This one is neglected in this tool and set to 1. This is how GNOM does it also.
The PDDF is $\gamma_0(r)$ and r is the distance (in A).
This tool uses two different methods to achieve its goal – Regularization (by Pete Jemian) and Moore's method (Moore, P. B. (1980). J. Appl. Cryst. 13, 168–175.). Both seem to have some advantages and disadvantages. The tool was tested against generally accepted GNOM by D. I. Svergun (EMBL). Test cases available to me yield same results and GNOM.

NOTE: I have observed significant changes on the calculated PDDF with changes with maximum dimension assumed and with errors scaling. My observations are noted in text further below.

Comment on graph formatting… This tool uses same default font and font size setting as other tools. You can change font sizes and used font for legend, text boxes and tags in "SAS" - "Other tools"-"Configure common items".

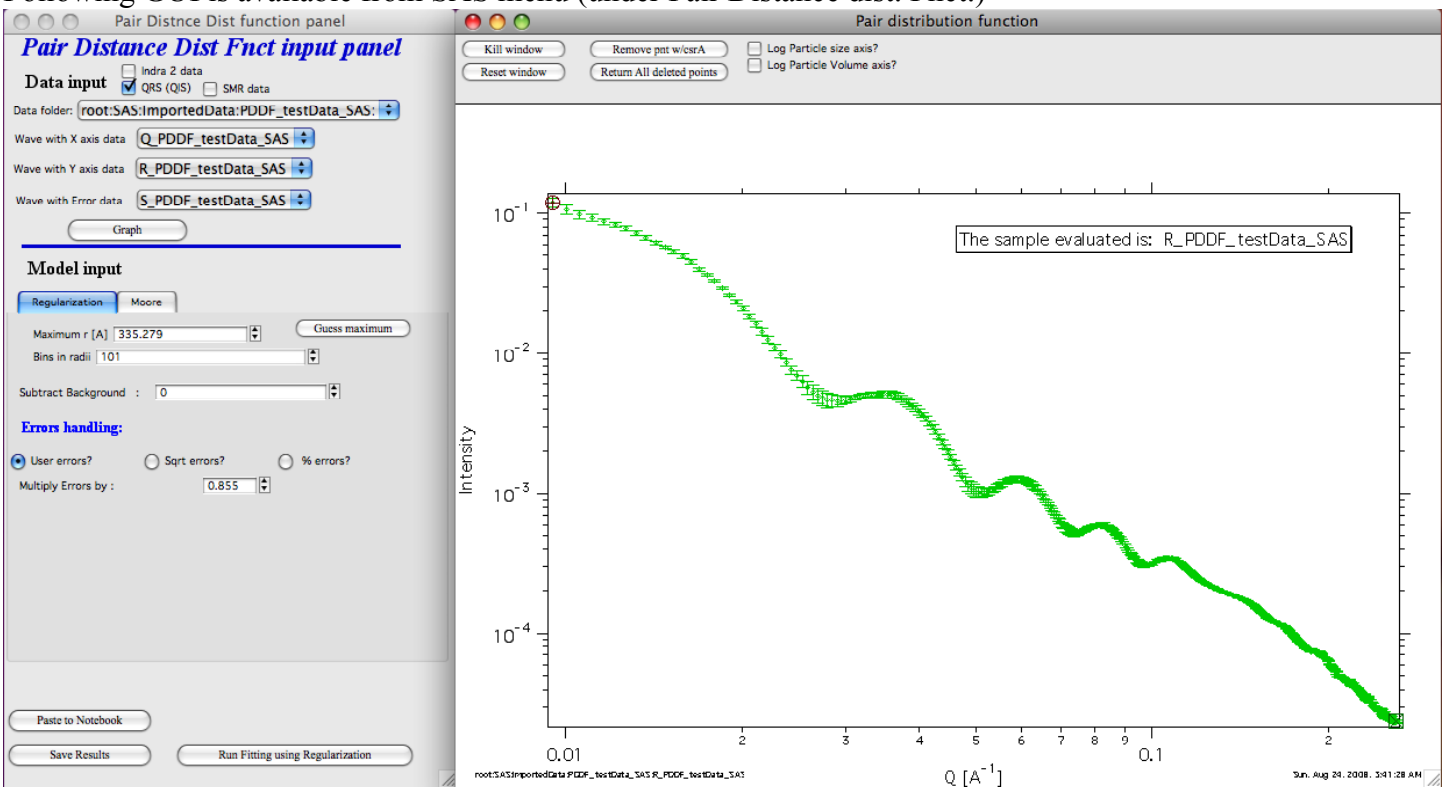
11.2 Use of the tool

Test data for this tool are included in the Irena folder distributed in the zip file and should be in your Igor Pro folder/User procedures/Irena



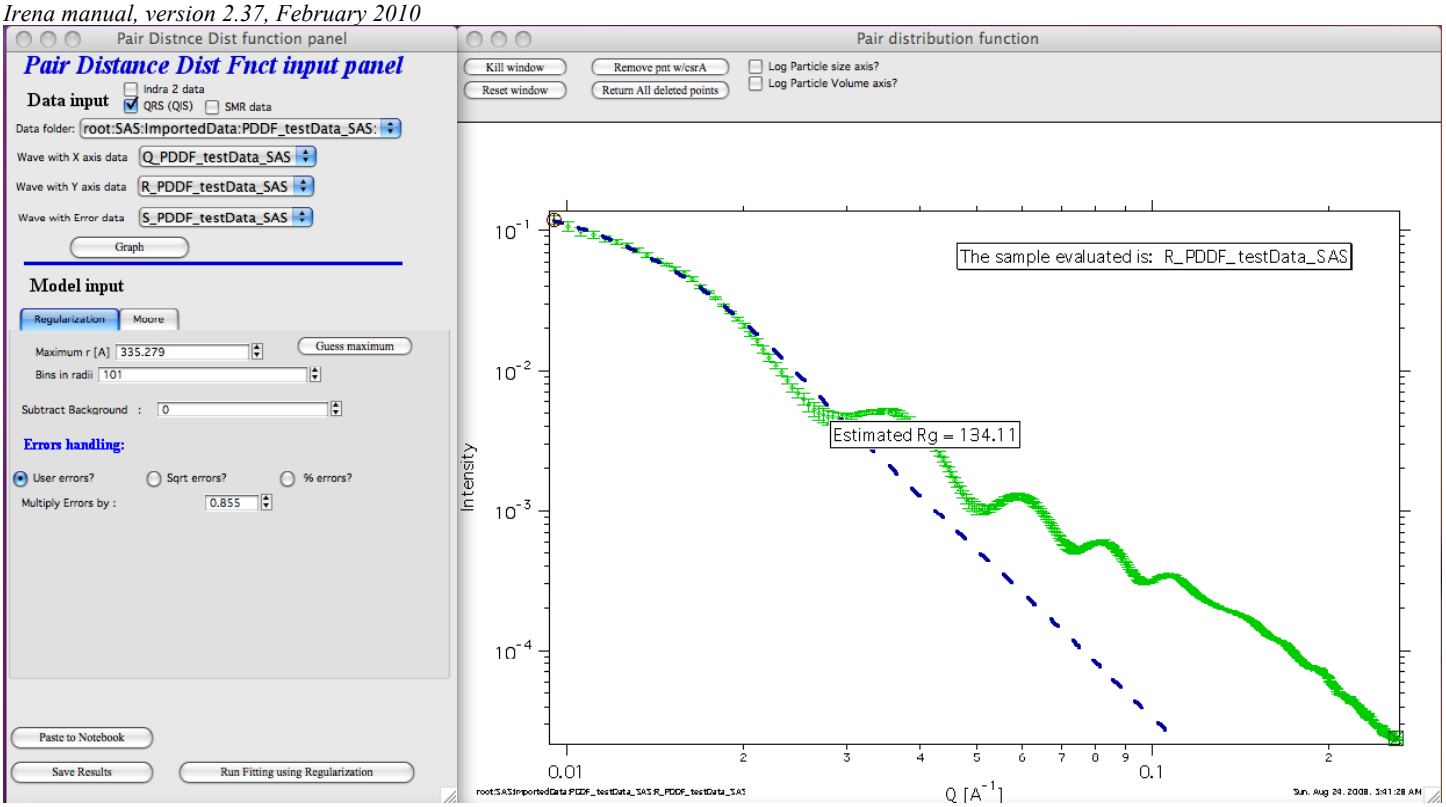
These data contain Q/Int/error and included is also GNOM generated PDDF in the similarly named file (see figure). These can be used to compare the results.
Load data in Igor as seen in the above figure and the follow next steps.

Following GUI is available from SAS menu (under Pair Distance dist. Fnct.)



In this GUI I have already selected the test data and pushed button “Graph”. This created the input graph on the right hand side.

Model Input selection:
PDDF modeling requires few right choices… Here are some suggestions how to get the right values for analysis…
1. Maximum r. Generally this is maximum distance for p(r) (=PDDF) function. For relatively spherical particles it is close to 2*Rg, for less spherical particles can get larger, may be up to 4* Rg. It is important to guess large enough number, but not too large. To help, you can try using the button “Guess maximum”. In this case the code will attempt to fit one-level Unified fit to the data and provide guess for Rg. Maximum r is set to 2.5*Rg. Here is result in this case:



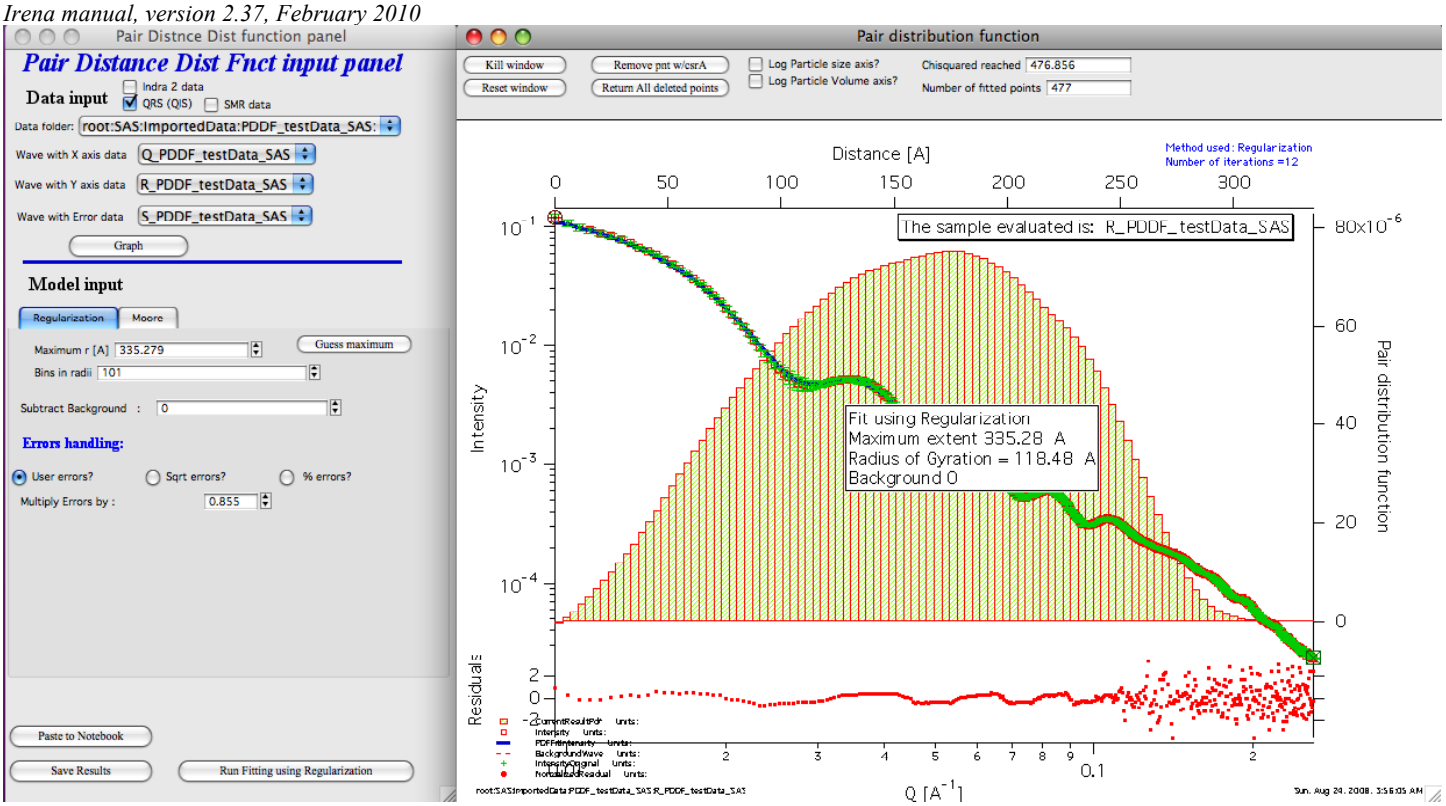
Note, this fit is not exciting, but the Rg is actually quite good, as you will see later…

2. Next one needs to choose number of bins. Too large number slows down calculations. I am not sure if higher numbers are of much use.

3. Subtract background – if there is some flat background in the data still left, one can subtract it here. Moore's technique can fit the background. Test data really do not have any background left.

4. Errors handling. There is no perfect selection here. One needs to play and get the right errors handling here. Many SAXS data reduction tools do not produce meaningful errors and each technique required somehow different error handling. "sqrt errors" are meaningful ONLY if the data are still in "counting" statistics. Rare case… However, there are some ideas about the right approach here:
Regularization
Start with higher error multiplier (for User errors of sqrt errors) and then try fitting with decreasing error multiplier. At some point the fit will look good – and when multiplier is decreased even more, the fit will start failing. Lowest multiplier when you can still get fit is probably close to right…
Moore technique
Uses least square fitting. I had better success with using fractional errors. Again, reduce errors to force good with within reasonable number of iterations.

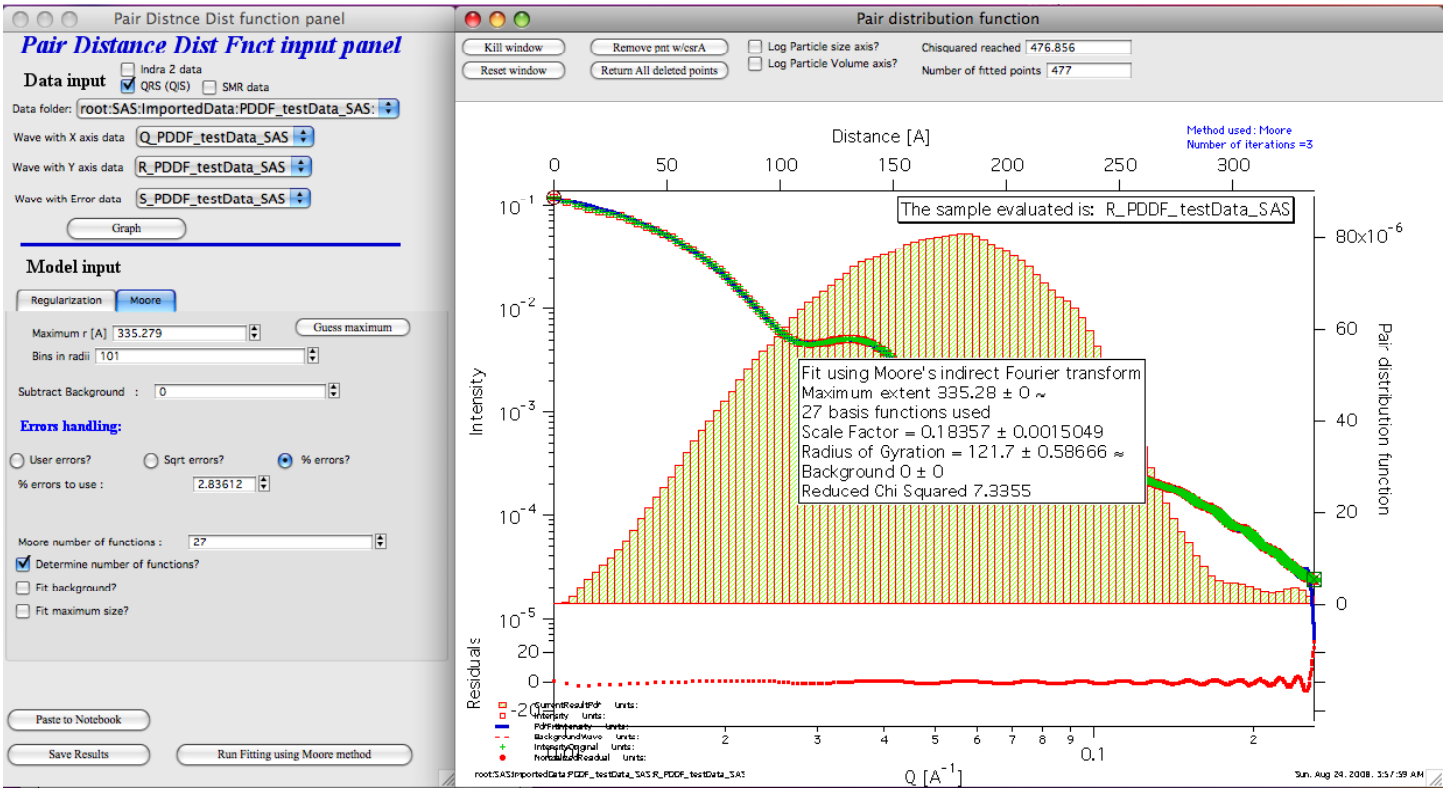
Regularization
There is nothing more needed, just select range of data to fit (probably whole range, but can be limited using cursors) and push fit button:



And here is result… One can see the PDDF, below graph are normalized residuals, provided is Rg and fit int eh graph.

Moore technique (indirect Fourier Transformation)

Select the tab with "Moore" and then see below:



Note, that one has more controls:

"Determine number of functions" – that is useful to make sure reasonable number of function is chosen… I suggest using it, unless you have reason not to.

"Fit background" – if there is flat background left in the data, you can try.
"Fit maximum size" – you can try, but in my experience resulting maximum size seems too low.

11.3 Semi-GNOM file and other output data methods

There are three buttons to use with three different methods to output data.

New in version 2.31 is output of Semi-GNOM ASCII file for use in other ATSAS packages. ATSAS is well known package of programs from Dmitri Svergun, http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html . GNOM is program which performs regularization method of PDDF analysis, same as PDDF in Irena package. Its output file is being used by all other ATSAS programs, such as DAMMIN etc. A user has requested that I provide method of outputting output file compatible with GNOM to use with results from Irena PDDF tool.
The GNOM file format does not seem to be publicly described and therefore, I had to reverse engineer which parts of the GNOM file are actually important for other programs and formatting of all different fields, as the formatting seems to be really unusual and obsolete.
The provided data format has been tested on DAMMIN PC version 5.3 and attempts to follow the GNOM file version 4.4 included as example with DAMMIN. I cannot guarantee any functionality. If you find case when it does not work, send me the Igor experiment and all other related details and I will try tooimprove the compatibility, if I can.

Note, not all parameters printed in the output file are meaningful for Irena PDDF tool. Some of them are there because they just seem to have to be there.

Here is snippet of the GNOM output file, red are my comments

```
*****************************
      ####   G N O M   ---   Version 4.4      ####Header, must be here

                              Thu Sep 25 08:44:00 2008          Date, meaningful
      ===   Run No   1   ===              meaningless
Run title:  root:SAS:ImportedData:lyzexp:R_lyzexp               Your data name, meaningful


*******   Input file(s) : R_lyzexp                      meaningful
      Condition P(rmin) = 0 is used.                    meaningless
      Condition P(rmax) = 0 is used.                    meaningless

     Highest ALPHA is found to be   1                   meaningless




       ####         Final results         ####                    meaningless

 Angular  range   :   from   0.0414   to   0.4984                        meaningful
 Real space range  :   from     0.00   to   50.00                    meaningful

 Current ALPHA       :  0.10E+01   Rg :   0.153E+02   I(0) :   0.655E+01       Alpha is meaningless, else is meaningful
 Real space range  :   from     0.00   to   50.00                    meaningful
```

S J EXP ERROR J REG I REG meaningful

0.0000E+01 0.6555E+01 meaningful
0.2299E-02 0.6552E+01
0.4598E-02 0.6544E+01
0.6897E-02 0.6530E+01
0.9197E-02 0.6512E+01
0.1150E-01 0.6488E+01
0.1379E-01 0.6459E+01
0.1609E-01 0.6424E+01
0.1839E-01 0.6385E+01
0.2069E-01 0.6341E+01
0.2299E-01 0.6291E+01
0.2529E-01 0.6237E+01
0.2759E-01 0.6179E+01
0.2989E-01 0.6116E+01
0.3219E-01 0.6048E+01
0.3449E-01 0.5977E+01
0.3679E-01 0.5901E+01
0.3909E-01 0.5822E+01
0.4138E-01 0.5904E+01 0.7150E-01 0.5739E+01 0.5739E+01 meaningful
0.4372E-01 0.5652E+01 0.7020E-01 0.5651E+01 0.5651E+01
0.4605E-01 0.5533E+01 0.6995E-01 0.5560E+01 0.5560E+01
….

Distance distribution function of particle meaningful

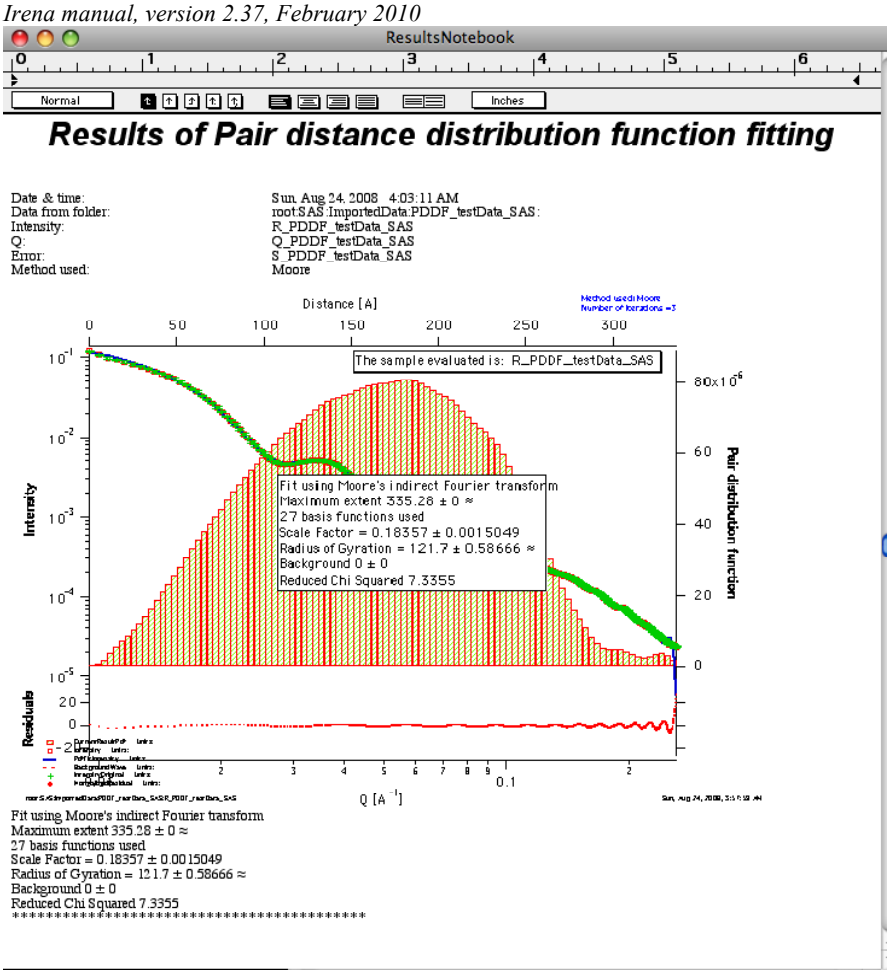
R P(R) ERROR meaningful

0.0000E+01 -0.5838E-03 0.5818E-04 meaningful
0.5000E+00 0.6171E-04 0.4782E-04
….

Reciprocal space: Rg = 15.252 , I(0) = 0.6555E+01 meaningful
Real space: Rg = 15.252 +- 0.000-00 I(0) = 0.6555E+01 +- 0.000E+00 meaningful, except for errors.

Other methods of saving data…

"Save results" copies wave with results into originating data folder. Copied are both model intensity and Q vector, as well as normalized residual. Also copied is PDDF and associated size wave. All of these waves have wave notes with all parameters and are recognized as results by Plotting tool, Data export tool and other Irena tools.
"Paste to Notebook" copies graph and somehow formatted summary of result into speciual notebook (created if necessary) for printing and future review.



You can access this notebook (if exists) from “SAS”-“Other tools”-“Show Results notebook” menu. You can save the notebook as RFT file, which then can be edited in any Word processor.

12. Fractal model

This model has been developed by Andrew J. Allen from NIST (Andrew.allen@nist.gov). The model allows to combine two volume and two mass fractals in much similar way as the Unified model does. The parameters from this model have advantage of being more "fractal-related" than the values from Unified. There is short pdf file included in the distribution which served as basis for my design of this tool. Note, that this tool is actually port of Andrews original Fortran code into Igor, my code was verified to give same results as this Fortran code.

Note, that this write up was written for studies of cement and therefore some of the terms are material-specifically called.

12.1 Model description

The model predicts Q^{-Dv} scattering (i.e. between Q^{-1} and Q^{-3}) for mass- or volume-fractals, and $Q^{-(6-Ds)}$ scattering (i.e. between Q^{-3} and Q^{-4}) for surface-fractals. In the model function for $d\Sigma/d\Omega$ as a function of Q, there are four components:

$d\Sigma/d\Omega$ = {VOLUME FRACTAL + SINGLE GLOBULE} TERM

+ SURFACE FRACTAL + FLAT BACKGROUND SCATTERING [1]

These components are incorporated into the full theoretical expression as follows:

$$\frac{d\Sigma}{d\Omega} = \phi_{CSH} V_p \left|\Delta\rho\right|^2 \left\{ \frac{\eta R_C^3}{\beta R_o^3} \left(\frac{\xi_V}{R_C} \right)^{Dv} \frac{\sin\left[(D_V - 1) \arctan\left(Q\xi_V \right) \right]}{(D_V - 1) Q\xi_V \left[1 + \left(Q\xi_V \right)^2 \right]^{(Dv-1)/2}} + \left(1 - \eta \right)^2 \right\} F^2(Q)$$

$$+ \frac{\pi \xi_{SS}^4 \left|\Delta\rho\right|^2 S_o \Gamma(5 - D_S) \sin\left[(3 - D_S) \arctan\left(Q\xi_{SS} \right) \right]}{\left[1 + \left(Q\xi_{SS} \right)^2 \right]^{(5-Ds)/2} Q\xi_{SS}} + \text{BACKGROUND} \qquad [2]$$

The first volume-fractal term contains ϕ_{CSH}, ξ_V, and the mean radius, R_o, and shape aspect ratio, β, of the building-block C-S-H gel globules in the volume-fractal phase, here assumed to be spheroids. It also contains a local volume fraction, η, and the mean correlation-hole radius, R_c: the mean nearest-neighbor separation of the gel-globule centers. R_c, assumed to be weighted over spheroid surface-contacts, is given by:

$$R_C = \frac{R_o \sqrt{2}}{\chi_s} \left\{ 1 + \left(\frac{2 + \beta^2}{3} \right) \chi_s^2 \right\}^{1/2} \qquad [3]$$

where:

$$\chi_s = (1/2\beta)\left\{1 + \left[\beta^2/\sqrt{1-\beta^2}\right]\ln\left(\left(1+\sqrt{1-\beta^2}\right)/\beta\right)\right\} \quad \text{for } \beta < 1, \qquad [4a]$$

and

$$\chi_s = (1/2\beta)\left\{1 + \left[\beta^2/\sqrt{\beta^2-1}\right]\arcsin\left(\sqrt{\beta^2-1}/\beta\right)\right\} \quad \text{for } \beta > 1 \qquad [4b]$$

In fitting the data, the need to incorporate R_c with η, and a well-defined single-globule term (in addition to the volume-fractal) in the first bracket of eq. [1], is strong evidence for a solid volume-fractal phase. A well-defined single-globule term arises because, unlike the case of fractal pores in clays and porous rocks, nearest-neighbor solid particles cannot exist inside each other, i.e., their centers cannot approach, on average, to within R_c. This correlation-hole effect means that, for length-scales of order R_o, the individual particles are seen as distinct objects, even when incorporated into an aggregated structure. For a spheroid of aspect ratio, β, the form-factor for a single globule, $F^2(Q)$, is given by:

$$F^2(Q) = \frac{\pi}{2}\left|\Delta\rho\right|^2 V_p^2 \left| \int_0^1 \frac{J_{3/2}\left(QR_o\left[1 + \left(\beta^2-1\right)X^2\right]^{1/2}\right)}{\left(QR_o\left[1 + \left(\beta^2-1\right)X^2\right]^{1/2}\right)^{3/2}} dX \right|^2 \qquad [5]$$

where $V_p = (4\beta\pi{R_o}^3/3)$, $J_{3/2}(x)$ denotes a Bessel function of order 3/2, and X is an orientational parameter, here integrated over all orientations of the spheroid with respect to Q. Use of a mildly spheroidal globule shape avoids the pronounced Bessel function oscillations for spheres ($\beta = 1$), which can perturb the fit at high Q. Satisfactory fits are obtainable with both mildly oblate ($\beta = 0.5$) and mildly prolate ($\beta = 2$) aspect ratios, giving globule sizes equivalent to a 5 nm sphere for cement.

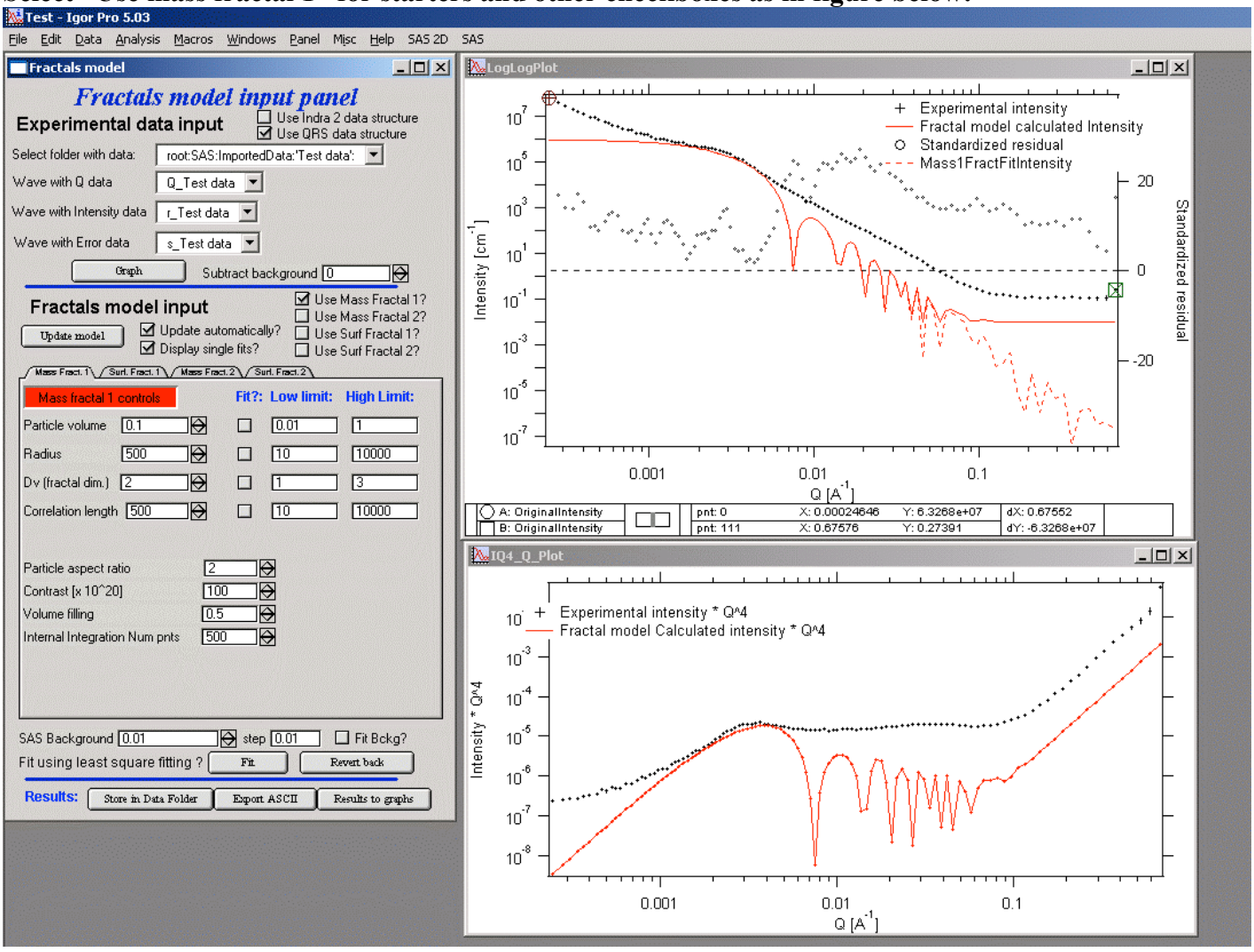
The surface fractal term in eq. [2] includes ξ_s, the mean upper limit of surface-fractal behavior at which the measured smooth surface area per unit sample volume is S_o. (The term, $\Gamma(5\text{-}Ds)$ is a mathematical gamma function.) The BACKGROUND term refers to the incoherent flat background scattering, and it is usually subtracted out of both data and fits for convenience.

12.2 Use

I do not have included real fractal data, but for purpose of GUI description and function description, the included data should be sufficient.
Start the tool from SAS menu under "Fractal model". GUI panel similar to all other tools appears, select "Use QRS data structure" and pick the data set available. The push "Graph" button to create graphs.

Note, that the "Subtract background" variable next to data selection allows to subtract known FIXED large background. The "SAS Background" at the bottom is similar term, but this one can be fitted during the fitting routine.

Select "Use mass fractal 1" for starters and other checkboxes as in figure below:

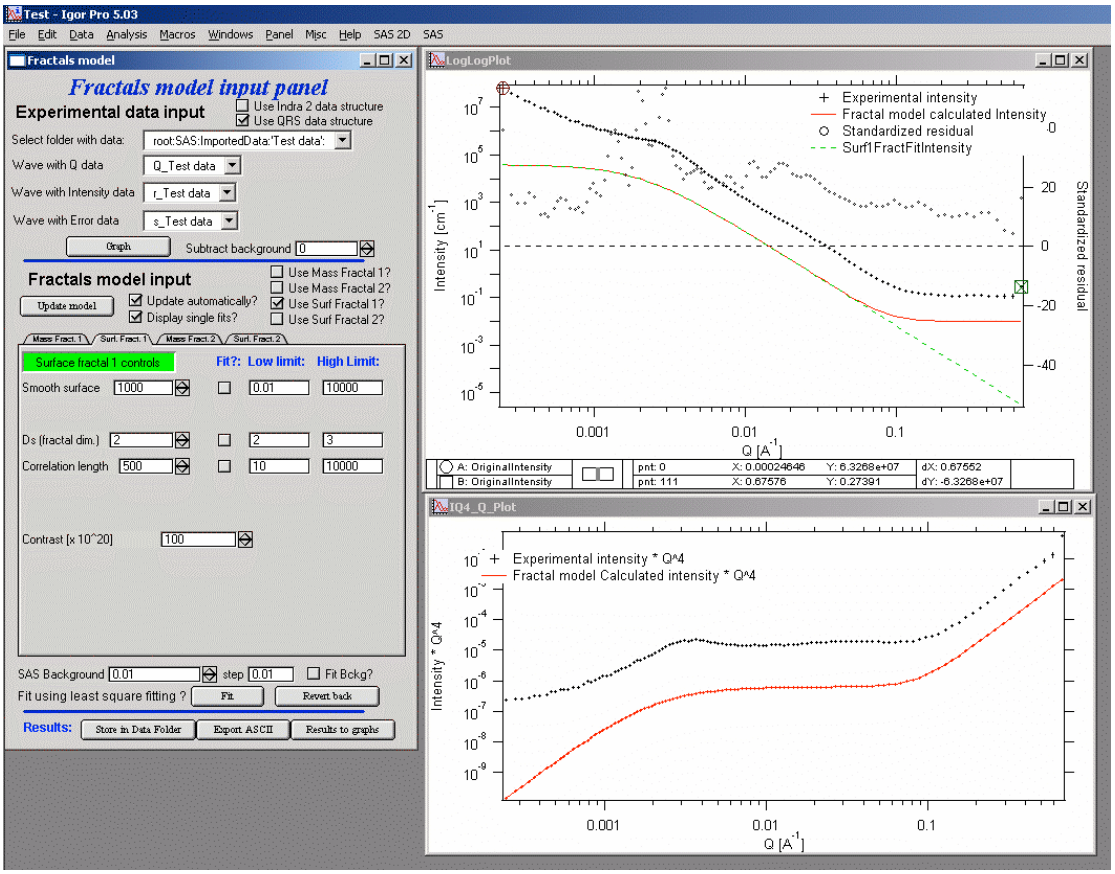


Note, that you can combine ANY combination of the two mass fractals and two surface fractals.

Comments on Mass fractal parameters:
Most parameters should be closely related to the ones mentioned above in description of the method.
Particle volume – volume of particles
Particle radius – size of the particle
Dv - fractal dimension
Correlation length – distance between the particles

Particle aspect ratio – 1 if particles are basically spheres, larger than 1 – elongated particles, lower than 1 prolated particles. Particles are always spheroids.
Contrast – contrast…
Volume filling – see above
Internal integration Num pnts – internal parameter. Number of point in the numerical integral which I use to calculate orientational average of the particle form factor. Small number of points (especially at high aspect ratios) can cause artifacts. Large number of points increases significantly calculation time. My suggestion is to lower the number of points to find a good starting conditions and for final fitting may be increase, or to recalculate for testing results with higher (double) number of points at the end – if no change is observed, the number of points is selected correctly.
Suggestions: check solution for particle aspect ratio 2 and 0.5, keep integral integration num of point reasonably high (over 100 for sure, likely around 500) and change it only if you seem to see artifacts. Keep volume filling between about 0.4 and 0.6.

Now select "Use Surf Fractal 1" and deselect the mass fractal:



Comments on surface fractal parameters:

Again, for meaning check the description above.
Smooth surface – limits of smooth surface as described above
Ds – fractal dimension
Correlation length – correlation length as described in the theory
Contrast - contrast…

Method of finding the solution is same as with Unified fit – first manually find good starting conditions and then select appropriate range of data with cursors and use fitting (select appropriate parameters to fit) to optimize data using least square fitting…

13. Analytical models (Debye-Bueche and Treubner-Strey)

This tool provides GUI for two different models:

Debye-Bueche model for modeling structural in homogeneities in the gels.

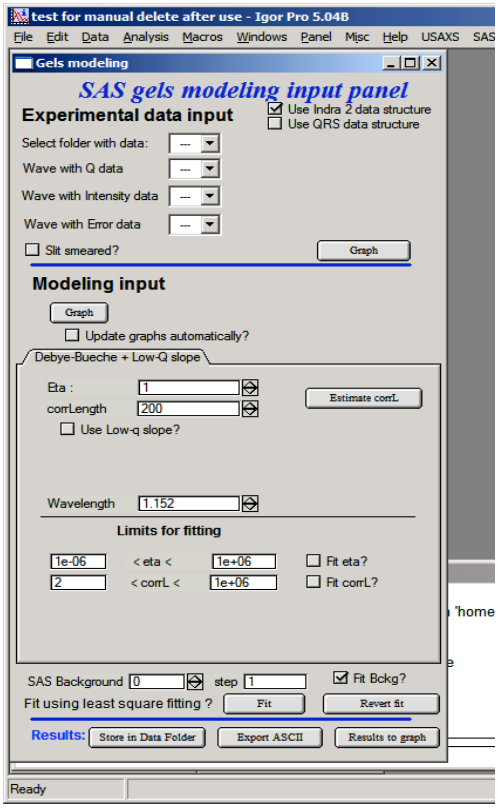
Treubner-Strey model for modeling of small-angle diffraction

Both models can be combined with low-Q Single Unified level. The controls have now three tabs – one for Unified level, one for Bebye-Bueche and one for Treubner-Strey. It is possible to combine them together, but it is not likely physically meaningful.

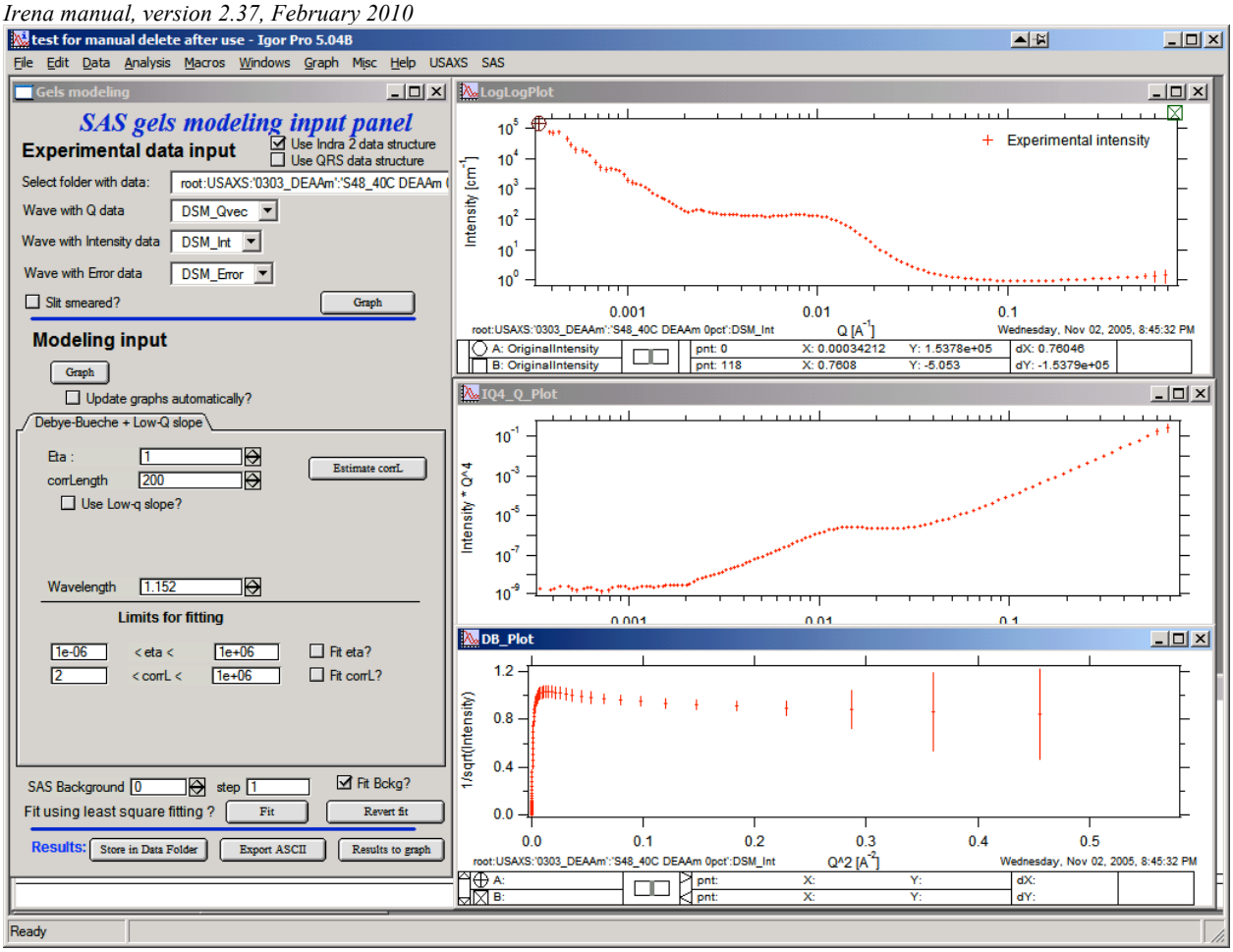
For explanation of the Unified level control, please see Unified fit.

Debye-Bueche model for gels

The theory is implemented in following form:
$I(q) = (4\pi K\varepsilon^2 corrL^2)/(1+q^2 corrL^2)^2$
where $K = 8\pi^2\lambda^{-4}$
Parameters of the gel are then the corrL – correlation length and ε. The model also allows low-q power law to be fitted and subtracted from data as well as flat SAS background. The low-q power law slope has 2 parameters (slope and prefactor) and background has one. All can be fitted.
This is the main screen:

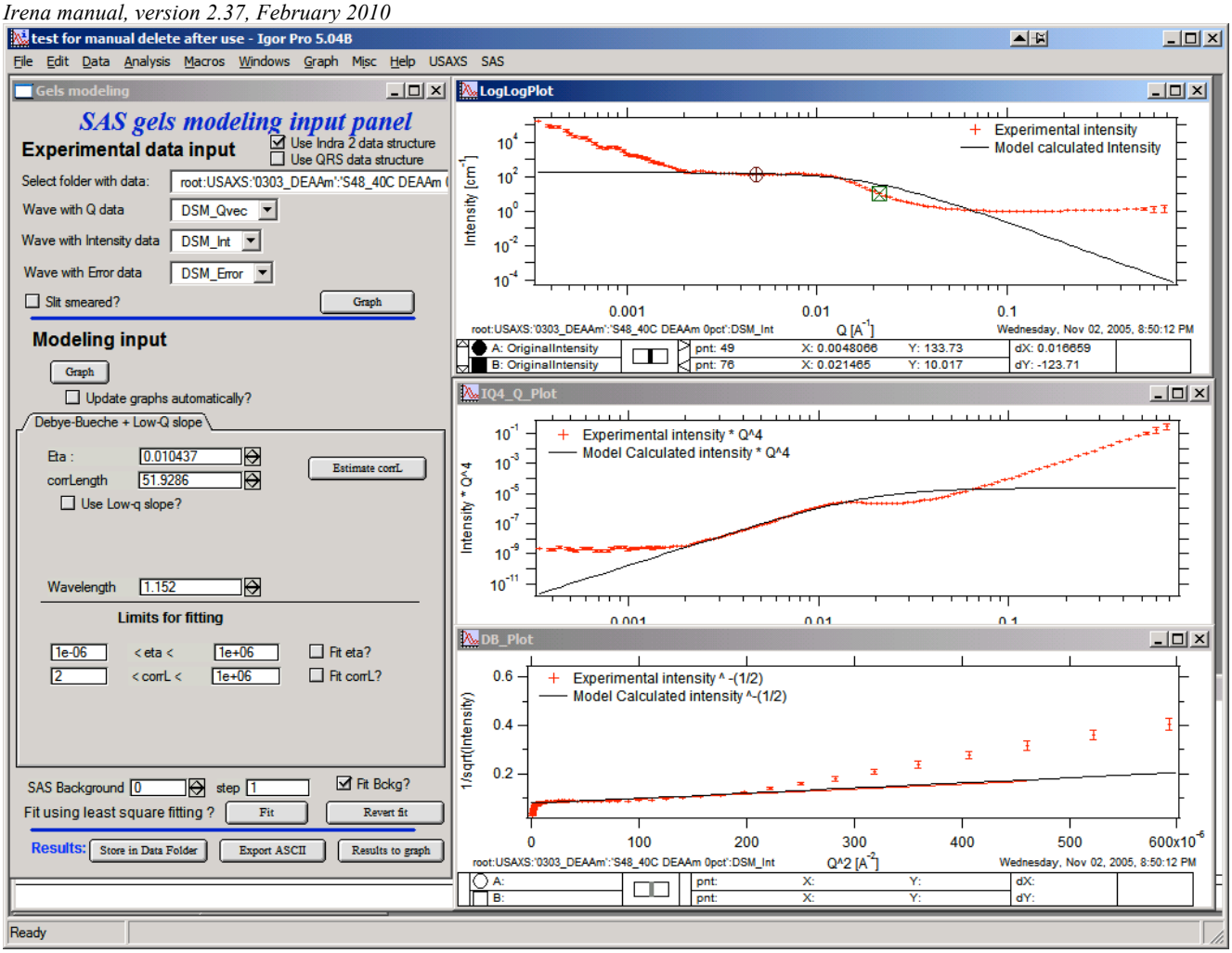


Data can be selected at the top part – as usually, one can use either pin-hole type data (desmeared for USAXS instrument) or slit smeared data. Results are the same, the model is slit smeared with slit length if slit smeared data are used.

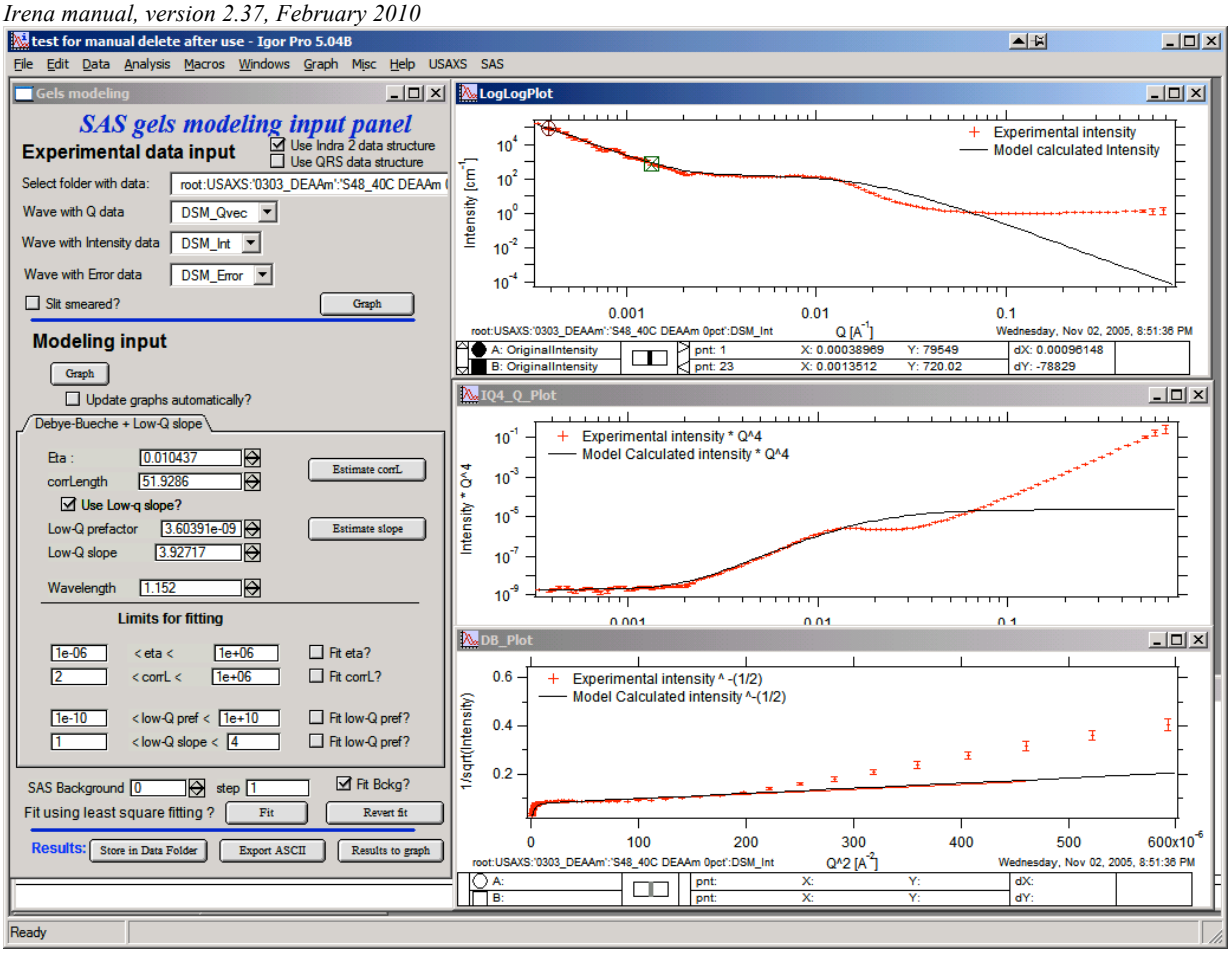


This is how the screen looks like with data selected. Note three graphs:
Top is log-log, middle is I * q^4 vs q, and bottom is 1/sqrt(Intensity) vs q^2. Data selection for fitting purposes is in the top graph…The other two are only for informational purposes.

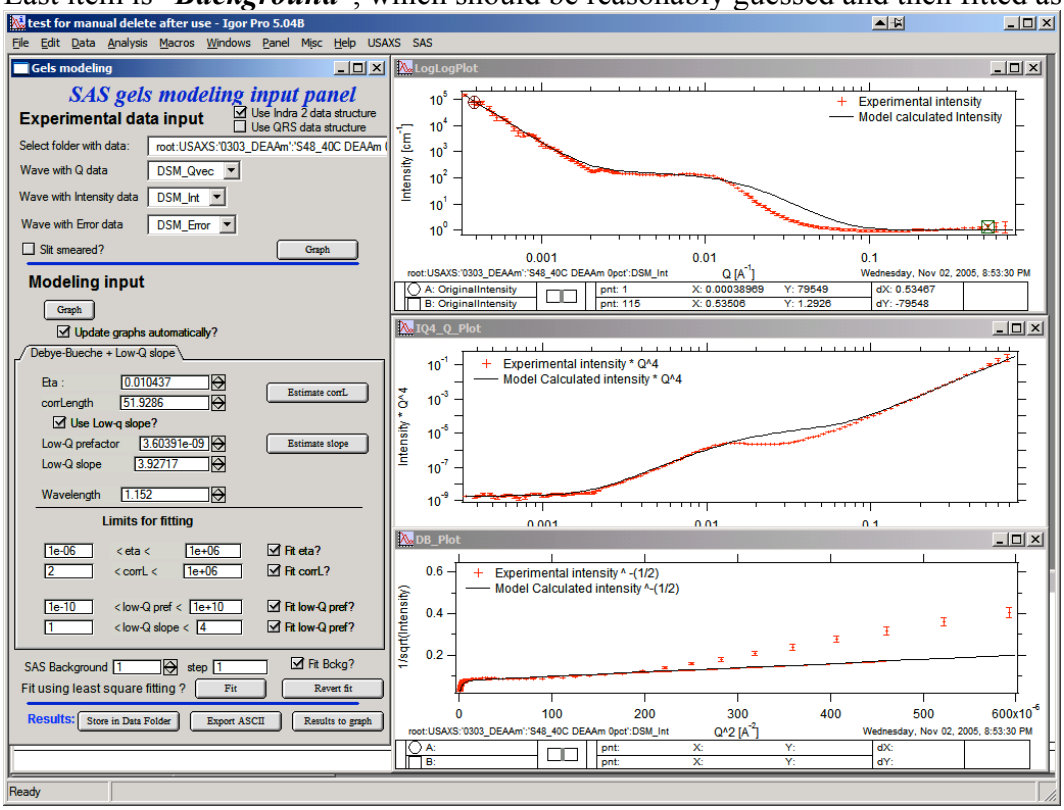
Controls:
Top button "***Graph***" loads data into the tool and creates the graphs.
Lower Button "***Graph***" will calculate model and place result in the graphs.
"***Update graphs automatically***" will recalculate model after every change of any parameter in this tool. Useful on fast machines.
Eta and ***corrLength*** – model parameters. Can be estimated using the button "Estimate" if the knee area is selected first in the top graph:



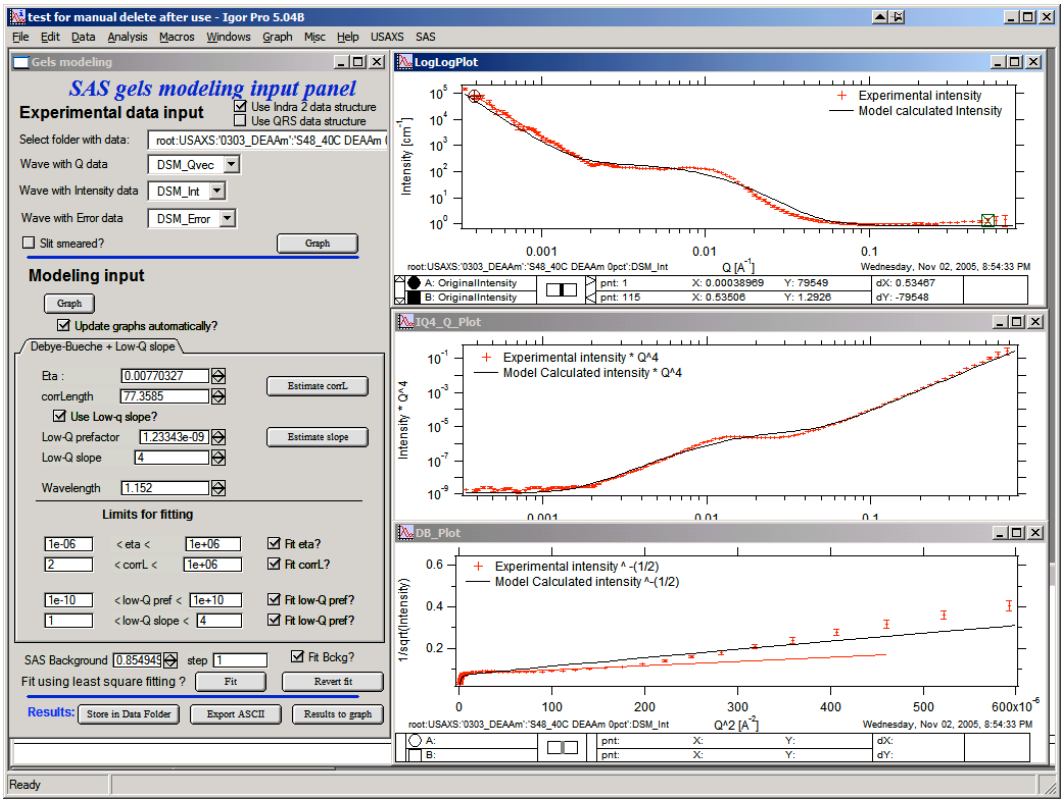
Checkbox "***Use low-q slope***" will enable controls for low-q power law slope. One can again select range of data where the power law dominates and Estimate slope with the button.



Limits for fitting should be set, if needed, to sensible numbers. The checkboxes with "***Fit*** …" allow selection of parameters which are going to be fitted using standard Igor least-squares fit.
Last item is "***Background***", which should be reasonably guessed and then fitted as one of the parameters:

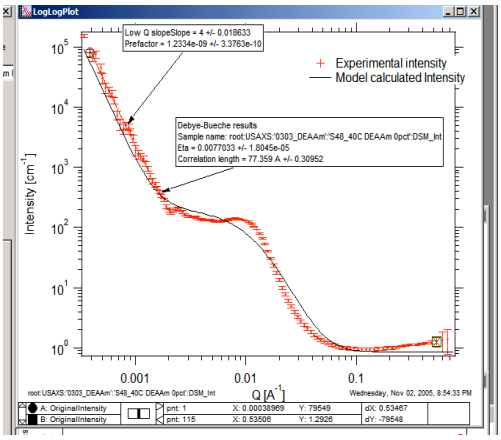


Now with good starting guesses one can fit the model – using the "Fit button"



This is the best fit this model does to these data (note the misfit, this is not probably the best model…).

Buttons:
Revert fit – use to reset the last set of parameters after bad fit which "lost it's way"…
Store in Data folder will save model data (waves with wave notes) for further use. It copies them into folder, where the data came from. Can be plotted, exported, reloaded in this tool, and mined for numbers later.
Export ASCII will export model as ASCII from Igor.
Results to Graph will paste results into graph for better view:



Treubner-Strey for small-angle diffraction

Treubner-Strey model follows the publications : Teubner, M; Strey, R. J. Chem. Phys., 1987, 87, 3195 and Schubert, K-V.; Strey, R.; Kline, S. R.; and E. W. Kaler J. Chem. Phys., 1994, 101, 5343.

The code is adopted form NIST SANS package. The formulas are:

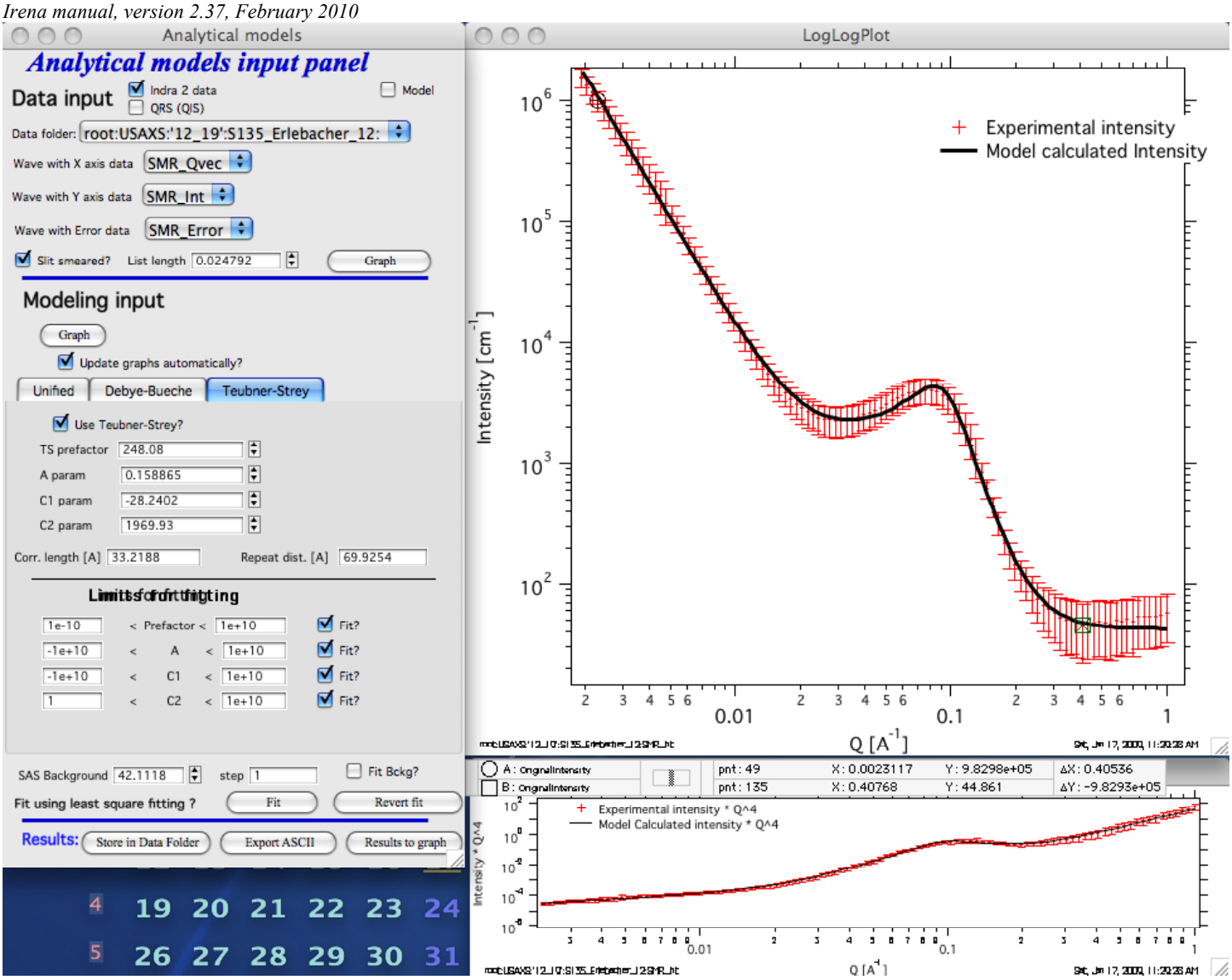
$$I(Q) = TS \frac{1}{A + C_1 Q^2 + C_2 Q^4}$$

Where A, C_1 and C_2 are parameters from the theory and TS is scaling factor.
Correlation length °ξ and repeat distance (d) are:

$$\xi = \left[\frac{1}{2} \left(\frac{A}{C_2} \right)^{1/2} + \frac{C_1}{4C_2} \right]^{-1/2}$$

$$\frac{d}{2\pi} = \left[\frac{1}{2} \left(\frac{A}{C_2} \right)^{1/2} - \frac{C_1}{4C_2} \right]^{-1/2}$$

Example of the GUI with results:
Note, that only the parameters TS, A, C_1, and C_2 are user controlled. Parameter TS is added scaling factor, as there does not seem to be other way to scale the model to data.



This is fitting to slit-smeared data for which Treubner-Strey model is the appropriate model to use.

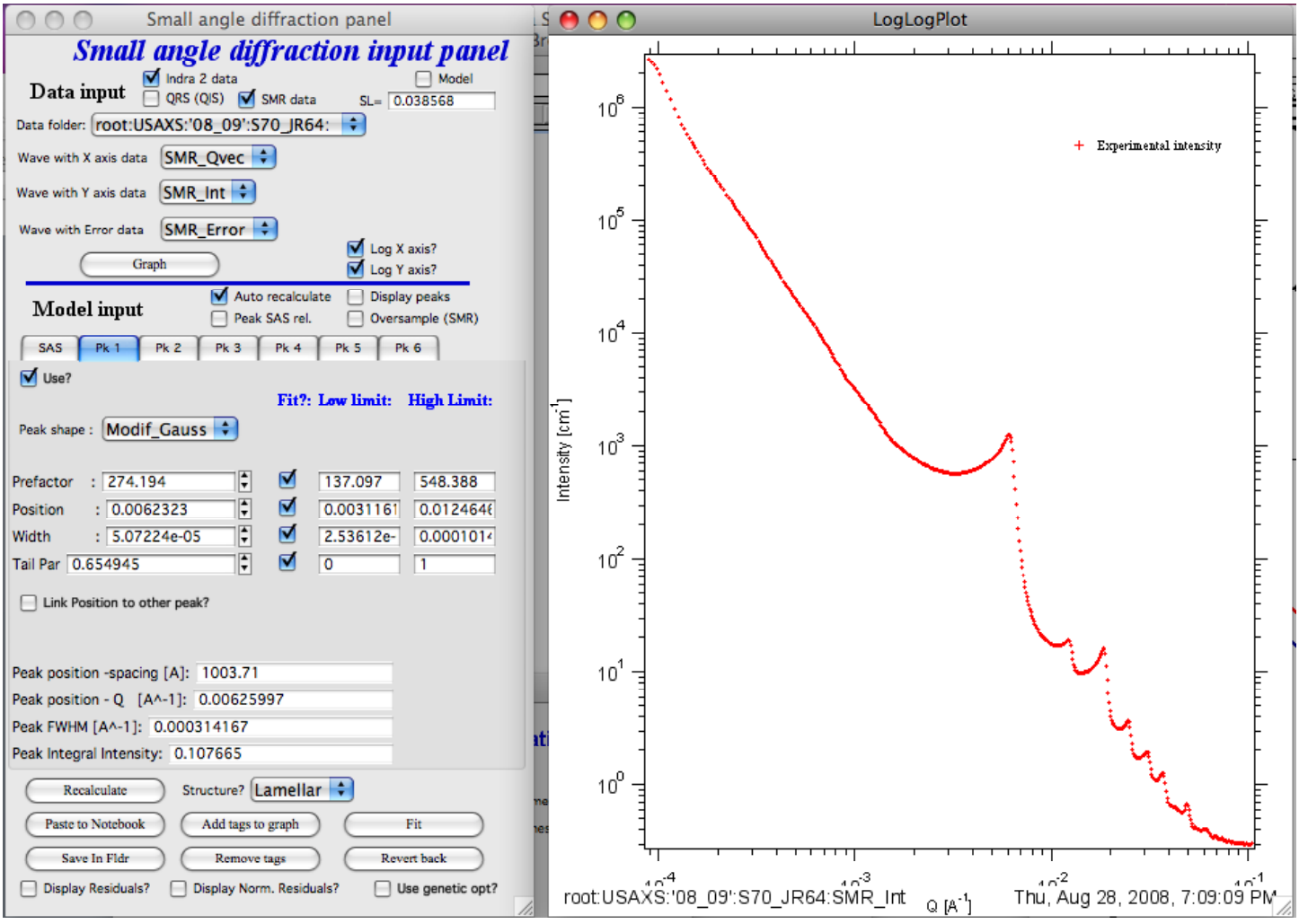
14 Small-angle diffraction tool

Small-angle diffraction tool models data using :
Flat background
Power law scatter ($I = G * Q^P$)
Up to 6 peaks

Each peak can have one of three different shapes – Guass, Lorenz, Pseudo-Voigt, Gumbel, Pearson-VII, and modified Gauss. Some first have 3 parameters – prefactor (~intensity), position (NOTE: using Q units) and width (in Q units). Some have one more parameter which controls the tail height or some other shape features. Note, that for Pseudo-Voigt when eta = 0 the shape is Gauss and eta=1 the shape is Lorenzian.

The tool will manage slit smeared data (USAXS data). There are few more details *very* important for slit smeared data:
It is very useful to use experimental data which extend significantly beyond to slit length. If the data to less than slit length are used, it is important to model peaks which extend to Q positions smaller than slit length. If you see ripples (caused by slit smearing very narrow peaks), you can use "Oversample" checkbox – but that will increase the calculation time by about 5x.

Use of the tool:
Select "Small-angle diffraction" from the menu



Select Data in the data selection controls and click graph button… Data are graphed.

Function of controls:

"auto recalculate" will cause data to be recalculated after most parameter changes. If calculations take long time, you may want to uncheck this and recalculate data using button "Recalculate".
VERY IMPORTANT
"Peak SAS rel." – this is very important checkbox. In case this checkbox is NOT selected, the following is the formula to calculate intensity:

$$I(Q) = I_{Unified}(Q) + \sum_i I_{Unified}(Q) K_i F_i(Q)$$

While when it is checked, then the formula is:

$$I(Q) = I_{Unified}(Q) + \sum_{peaks} K_i F_i(Q)$$

Where K_i is scaling factor for each diffraction peak.

Where Ψ(Q) is function of the three or four peak parameters – scaling factor, peak position, width, and for some also "tail" parameter. The exact formulas vary depending on peak profile selected.
What does this mean? If the checkbox is NOT selected, the calculation is based on assumption, that the SAS scattering and diffraction peaks are from one population and loosely one can see it as F(Q)*S(Q) assumption in small-angle scattering.
If the checkbox IS selected, the assumption is loosely that the peaks are independent of small-angle scattering and are produced by some other features than what produces the SAS itself.
I suspect, that right selection is based on experience and what really fits right. Note, that the parameters are always evaluated for Ψ(Q) only… This is *VERY* important to understand and if you see cases, when these assumptions are wrong, please, let me know…

Following are formulas for peak profiles Ψ(x) used for the peak profiles:

1. Gaussian Function

$$\Psi(x) = Me^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (1)$$

where σ is the Gaussian width, and μ is the center of the peak, and M is scaling factor.

2. Modified Gaussian Function

$$\Psi(x) = Me^{-\frac{(x-\mu)^d}{2\sigma^d}} \qquad (2)$$

where $d \geq 1$ is the exponent that decides the falloff rate.

3. Lorentzian Function

$$\Psi(x) = M\frac{a}{\pi(a^2 + (x-\mu)^2)} \qquad (3)$$

where *a* is the Lorentzian width.

4. Pseudo-Voigt Function

$$\Psi(x) = M(\eta \frac{1}{(1+x^2)} + (1-\eta)e^{-(\ln 2)x^2}) \quad (4)$$

where $x = \frac{2(x - x_0)}{w}$, x_0 is the peak center, w is the FWHM, and $0 \le \eta \le 1$ is a weight parameter.

5. Pearson type VII Function

$$\Psi(x) = M\left[1 + \frac{(x-\mu)^2}{ma^2}\right]^{-m} \quad (5)$$

where a is proportional to the FWHM, and m decides the rate at which the tail of the peak profile falls.

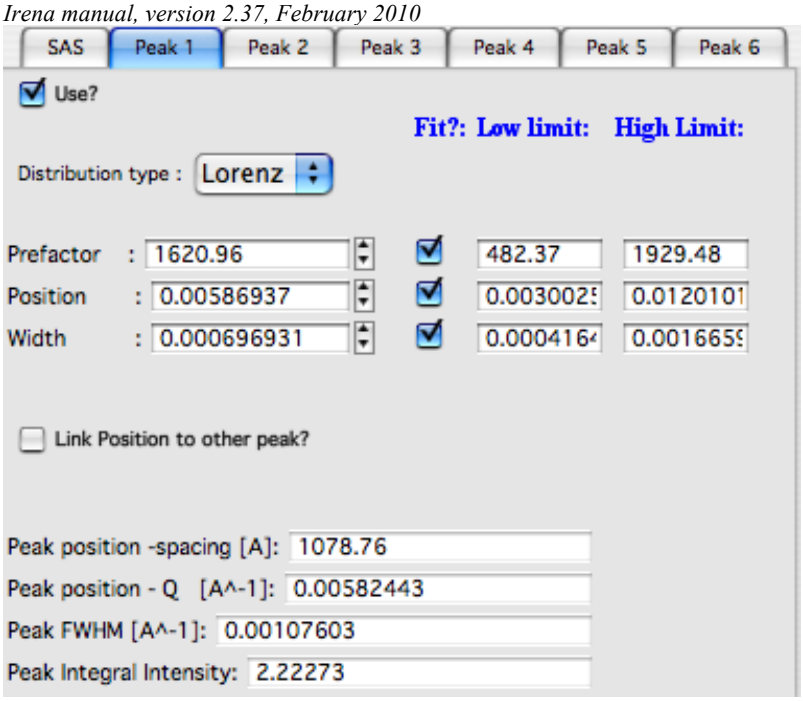
6. Gumbel Function

$$\Psi(x) = \frac{1}{\beta} e^{\frac{x-\mu}{\beta}} e^{-e^{\frac{x-\mu}{\beta}}} \quad (6)$$

where β is the width and μ is the center of the peak.

"Display peaks" will display individual peaks. Note, data for individual peaks are never smeared.
"Oversample" – for sit smeared data only. Will oversample Q range with 5x as many point to reduce artifacts caused by slit smearing very narrow peaks.
Tab SAS:
G – prefactor for power law slope
P – power law slope
Bckg – flat backgroud

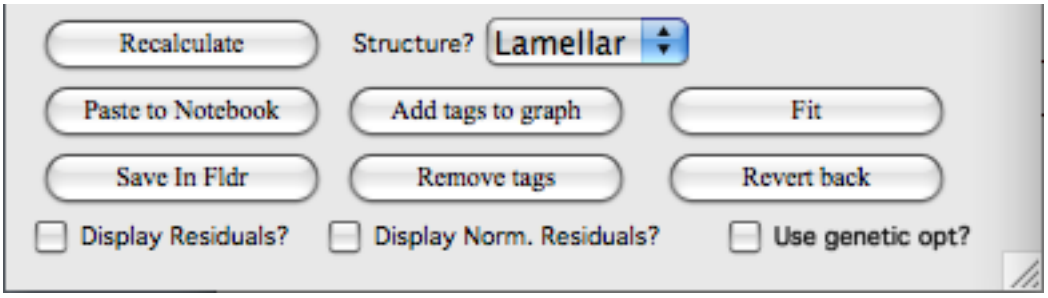
Tabs for Peaks:



"Use" – use the peak. No need to use peaks in order, can be mixed-and-matched
"Distribution type" – peak shape
"Prefactor" – scaling factor for the peaks (~hight)
"Position" – peak position in Q units
"width" – peak width in Q units
"Link Position to other peak?" – you can link peak position to position of another peak with scaling constant.

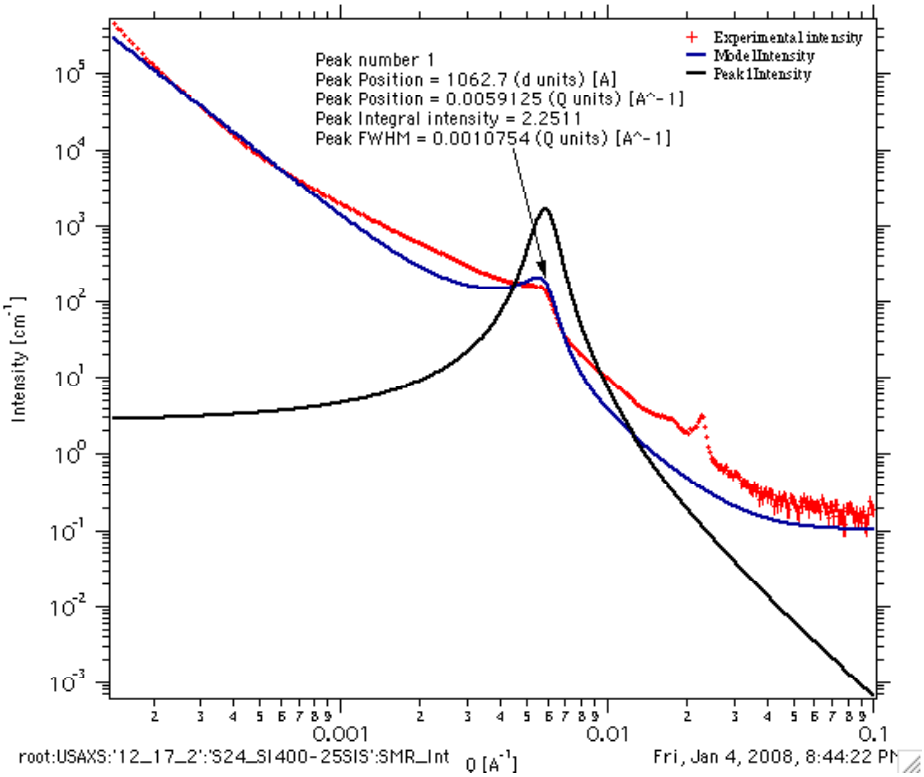
Lower set of parameters are peak parameters calculated numerically, so they may be slightly different than the numbers above.

Final controls:



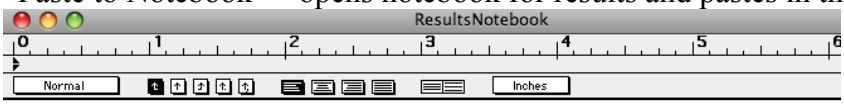
"Use genetic optimization?" – uses genetic optimization… Very slow fitting routine unlikely needed for this application. If needed, read explanation of the method in previous chapters.

"Fit" – fits
"Revert back" – reloads stored parameters from before fitting.
"Add tags to graph" – adds tags with parameters into the graph…
"Remove tags" – removes tags from the graph.
"Structure?" – sets ratios of positions for some known structures. Peak positions will be fixed with respect to Peak1. Note, user must set correct widths and prefactors for each peak manually…



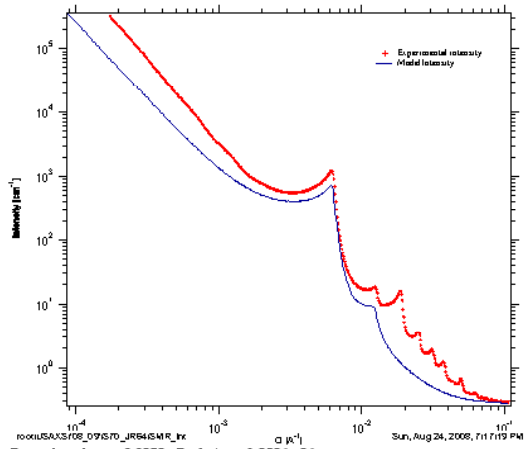
"Save in Fldr." Saves results (including peak profiles if selected) back into data folder.

"Paste to Notebook" – opens notebook for results and pastes in there graph and summary of results.



Results of Small-angle diffraction fitting

Date & time: Sun, Aug 24, 2008 8:35:23 PM
Data from folder: root:Packages:IrenaControlProcs:IR2D_ControlPanel:
Intensity: ModelInt
Q: ModelQ
Error: ModelError



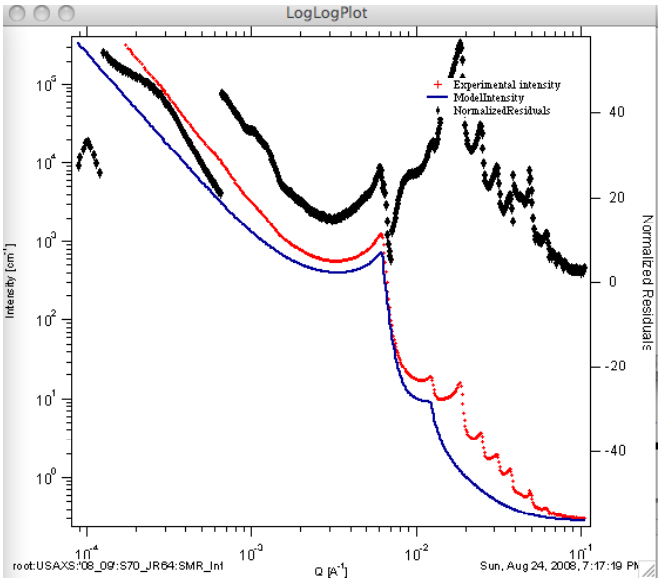
Power law slope = 3.3879 Prefactor = 2.9596e-06
Background = 0.2738

Peak number 1 used
Peak position (Q units) = 0.0063381 [A^-1] , D units = 991.34 [A]
Peak FWHM (Q units) = 0.00037309 [A^-1]
Peak Integral intensity = 0.11316

Peak number 2 used
Peak position (Q units) = 0.012553 [A^-1] , D units = 500.55 [A]
Peak FWHM (Q units) = 0.00089727 [A^-1]
Peak Integral intensity = 0.013875

****

"Recalculate" – forces model recalculation if use needs to do it.
You can attach also residuals or normalized residuals into the graph, see example below.



Useful comments:
Make sure the fitting parameters ranges are set appropriately. This is IMPORTANT and not obvious problem in fitting (experience speaks)… Results of fitting are also automatically recorded to into usual "SAS logbook" these tools keep… All is recorded there in more or less useful form. Your notes I keep for you....

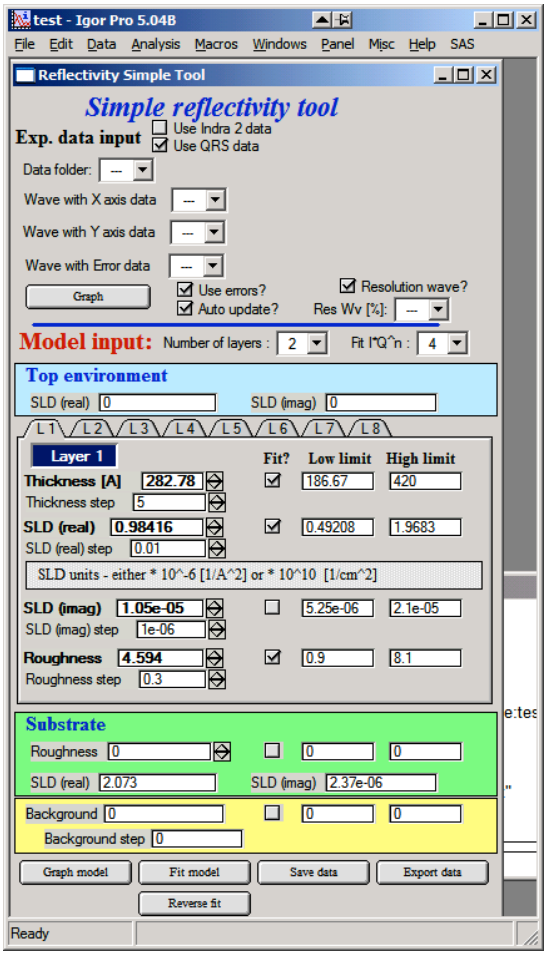
15. Reflectivity

This is relatively simple tool to model and fit X-ray and neutron reflectivity for up to 8 layers using recursive Parratt's code (L. G. Parratt, *Phys Rev*, **95**(2), 359-369, 1954), as implemented for example in older code called "Parratt". The code itself was provided by Andrew Nelson (Andrew_Nelson@users.sourceforge.net). I have implemented only the GUI in manner similar to the rest of the Irena code. I will provide basic support for this package. Please note, that Andy has made more complex and capable version of his package "Mottofit" available for free download to other Igor users through http://motofit.sourceforge.net/ under GNU license. If you need more complex fitting, than my code allows, please use Andy's powerful code. You may, however, have to learn little bit more of Igor.

Use of xop to speedup the calculations
Starting from version 2.20 (note: this version is ONLY 6.0 compatible) the code supports optional abeles.xop and GenCurveFit.xop. These are both optional – but the increase in speed (especially abeles.xop) is major (factor of 5-10x). These xops can be downloaded from Andrew Nelson's web site (listed above).
Note: If you have version 2.16 (last Igor 5 compatible version) note, that these xops will not be used. The code to use them was removed to make this release Igor 5.0 compatible.

15.1 Running the reflectivity

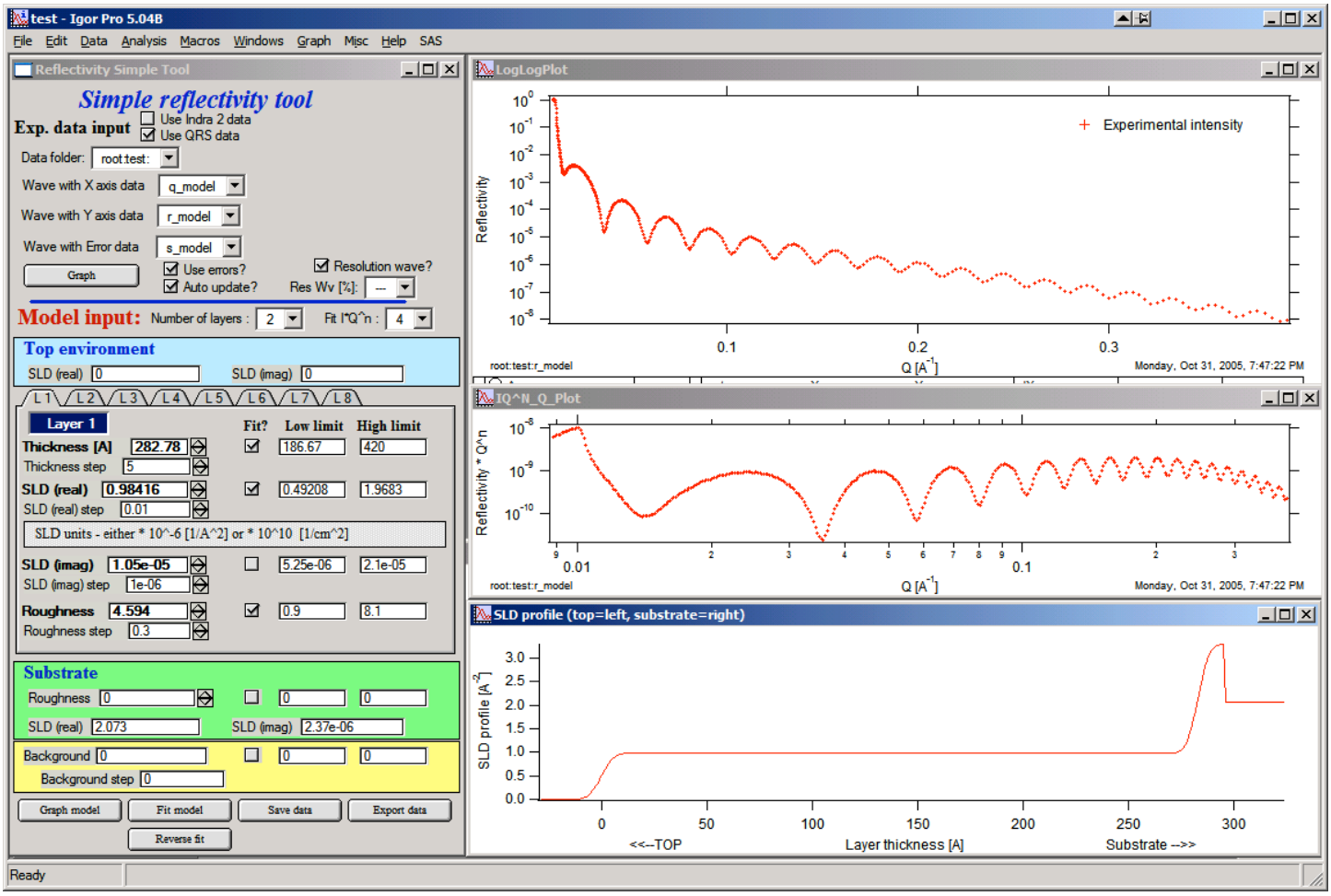
For testing purposes, I have included 3 column reflectivity data in Irena folder (…Wavemetrics/Igor Pro/User procedures/Irena) in the file called reflectivity.txt. Please load the data through Data loading tool as qrs data. The select "Reflectivity" from the SAS menu:



Following are the parts of the panel:

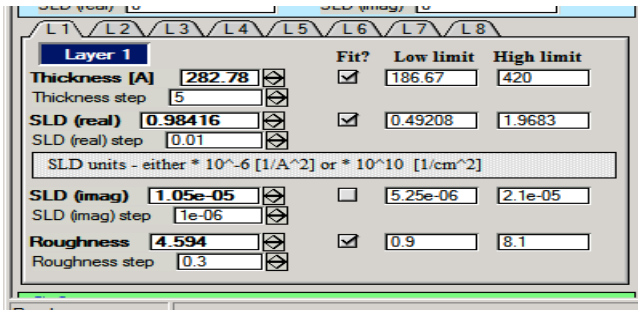
1. The top part – data selection part is similar to the other tools in Irena – select type of data, data folder with the data, and wave names containing q, reflectivity and error. Push Graph to generate graphs. If you have resolution as single value, uncheck the "Resolution wave" checkbox and insert value in the field, if you have resolution wave (q resolution), select it's wave name here. NOTE: for now this wave must be in the same folder as the data are. I will fix this in the future. Also, for now there are no checks on wave length, so make sure this wave has same number of points as data waves have.
2. Select number of layers, input scattering length density (SLD) for the top environment (usually air, so 0 is fine, but if this experiment was done for example under water, than this would be different).
   The tabs contain controls for each layer – thickness in A, SLD (real and imaginary) in units as displayed on the panel and roughness.
3. Substrate values – roughness and SLD
4. Measurement (aka flat) background.
5. Control buttons

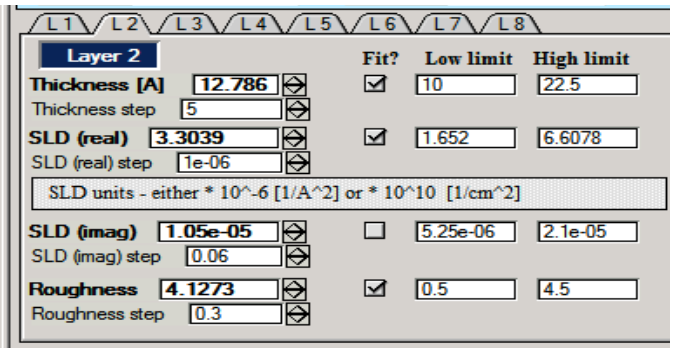
When data are selected, following graphs appear:



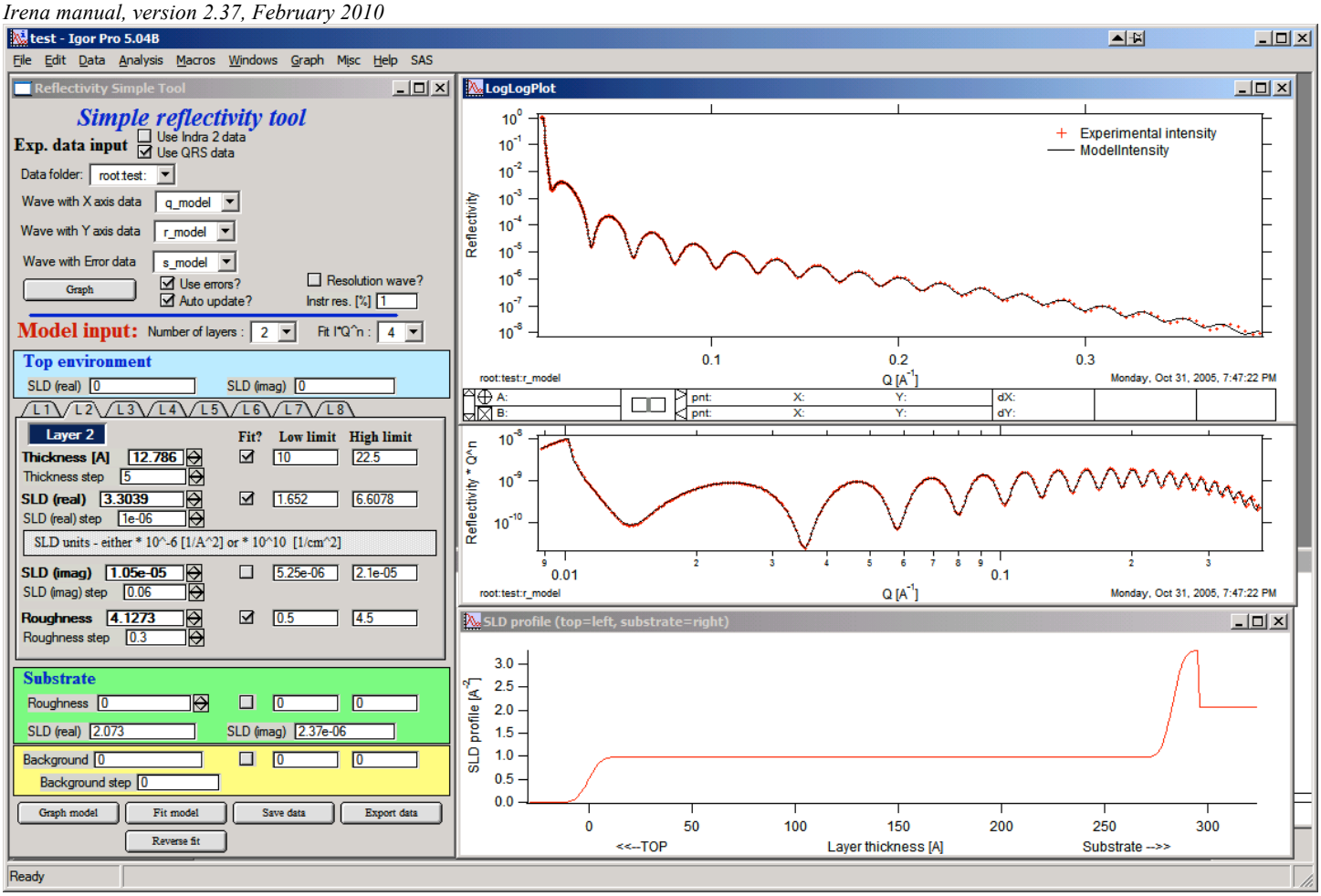
Note, that the top graph is log-log plot of reflectivity vs Q, medium is reflectivity * Q^n (n=0 to 4 as selected in the panel) and bottom is reflectivity profile. The fitting (see later) is done in the space reflectivity * Q^n to improve mathematical stability and convergence of the problem. The controls (selection of data range) however, MUST be done in the top graph (the log-log plot).

Now, I have very good values for these particular data (thanks to Dale Schaefer for providing the data and solution!!), see the tabs below:





Input these values and SLD for substrate of (real part) 2.073 and imaginary part 2.37e-6. The resolution is 1%, so uncheck the “resolution wave (if checked)” and input 1%.
Then push button “Graph model” you should get really good match to data:



I suggest you play now with parameters to find out, how sensitive the problem is.

Important comments

Resolution wave for now must contain % resolution for each point.
Fit model/reverse fit: Select range of data to fit in the top graph and push the Fit model button. Fitting is done in the Intensity * Q^n as selected in the panel. Use power of 4 is suggested, if lower values are used, the fitting tends to neglect the high-q data. If fit fails but reaches some solution, you can recover to previous data by pushing “reverse fit” button. Very handy…
Save data – copies model data into data folder so they can be used in the future. If you try to load data from folder containing already reflectivity data, you will have option to reload previous solution into the tool. This allows very quick regraphing of the stored solution.
Export data saves ASCII file outside Igor for use in other packages.

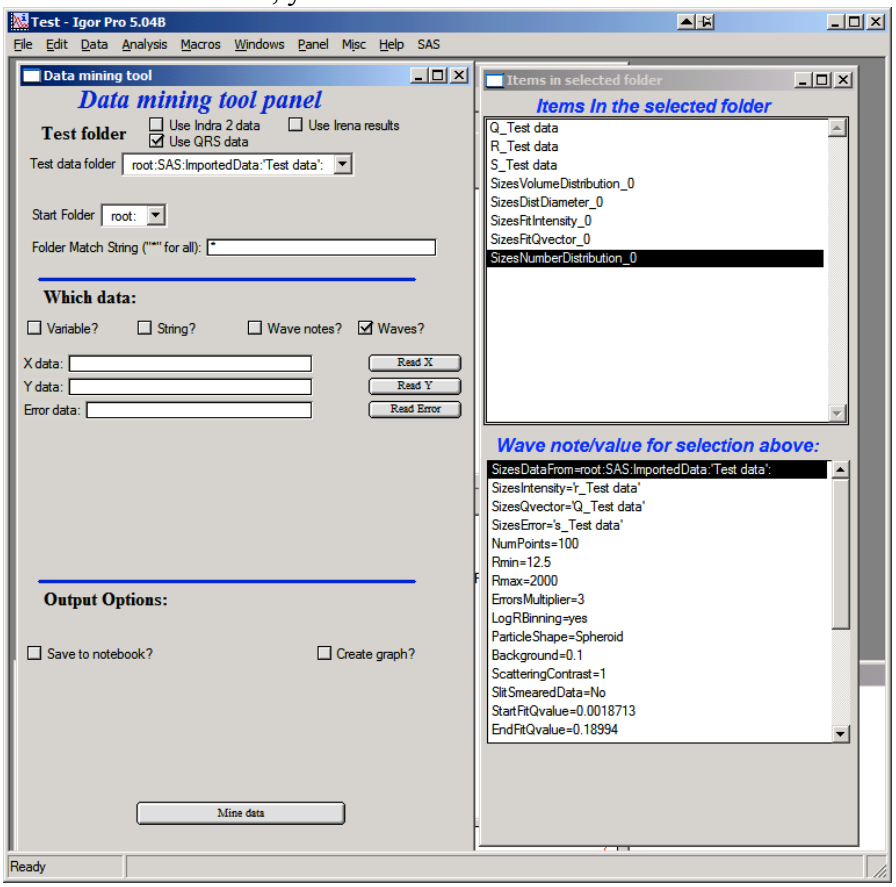
16. Data mining tool

This tool allows finding and tabulating various data from folders within Igor. At this moment it is just about functional and will be developed further.

The data can be stored in either variables, strings, wave notes or they can be waves with data themselves. Depending on input type, the data can be output into graphs or notebook (for waves), or in notebook or in new waves (for strings, variables and data from wave notes).

16.1 Use

There are few peculiarities, this tool behaves little bit different than the others… See below:
In the top part you need to select data type to be searched and example folder, in which are data types you want to "mine". Therefore, you need to be able to find at least one folder with data you are looking for.



When you select such folder, new panel with list of particular items from this folder appears. The top lists waves/strings/variables as appropriate; bottom lists the wave note (for waves) or the value of the string or variable.

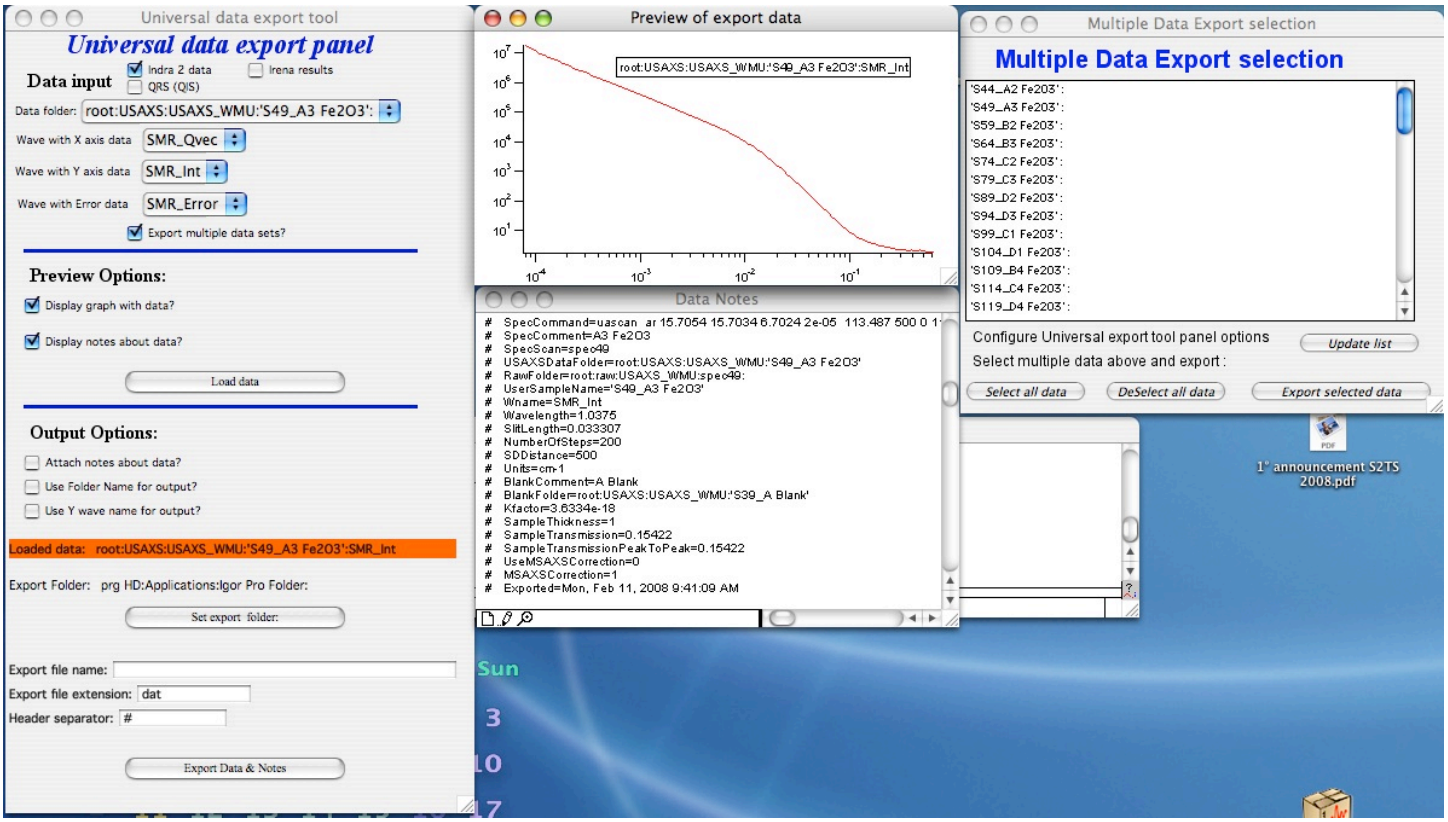
Mining the waves:
You can select (or input manually) wave names – one, two or three – in the fields. To select highlight the wave in the panel and then push button "Read X" (Y or Error). You can also use * as wildcard. Then you can either output into notebook (and print which folders contain – or do not contain these waves) or you can plot the waves into graph (plots X against Y) with few controls (colors, axis type).

Mining the strings/variables/wave notes
Select string or variable name in the list in the right panel and push button "add to list". The item will be added. To select wave note, select both wave and item in the wave note and pus button "add to list". Item will be added. At this time you have to Clear the whole list, if you want to remove something.

Output can be done to notebook (can be messy very fast…) or to new waves. Create new folder for the waves. After search, these waves will be created, if possible converted to number waves and new table will be created.

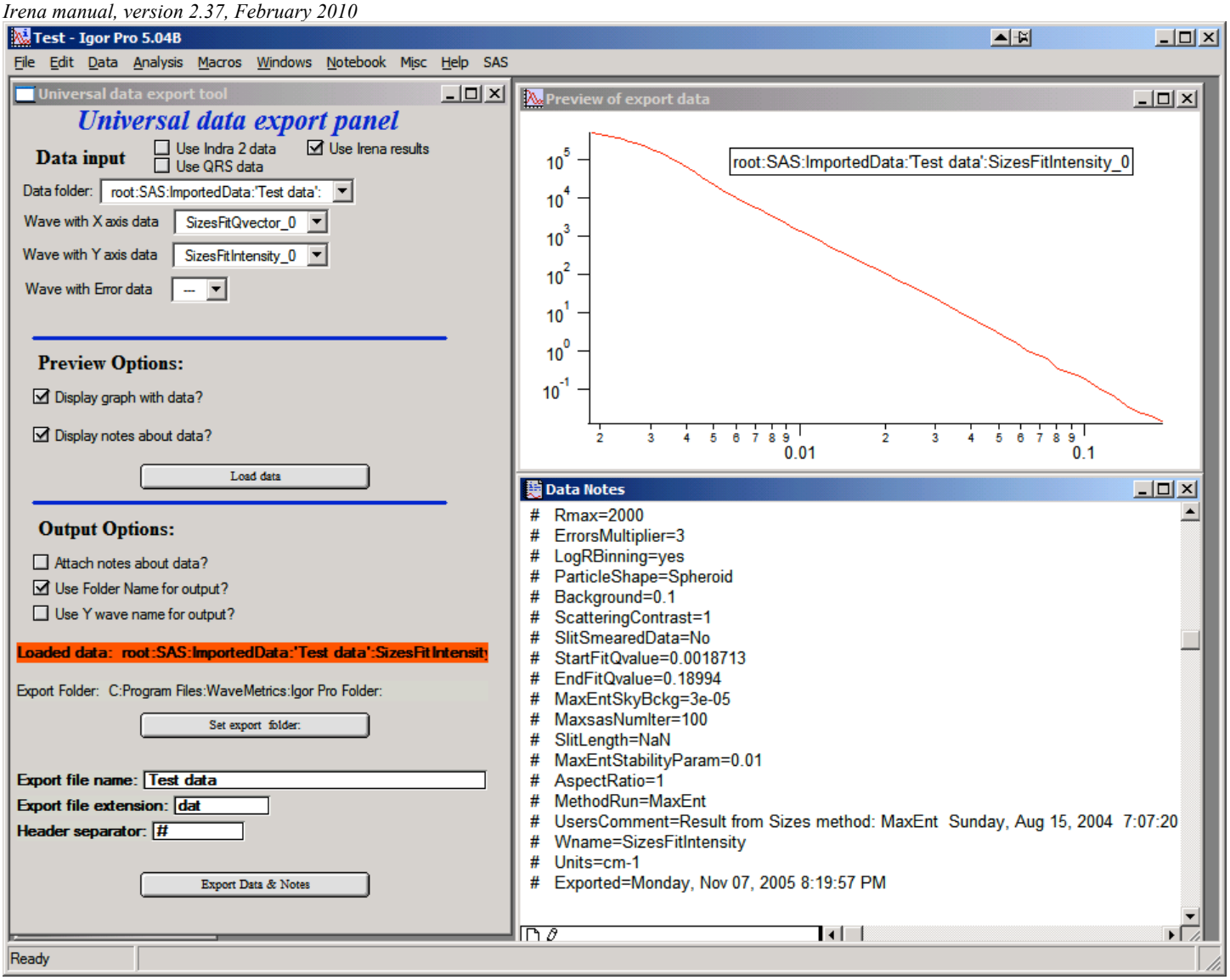
17. Universal data export tool

Irena produces large number of data sets, which can sometimes be exported when created, but more often not. While the most convenient use of the data is within Igor experiment by ploting or processing further, many users may want to use another programs. And then it is imperative to export them as ASCII files. This is the tool for that…



Above are standard set of controls. Select data in the top part of the panel. Than decide – do you want to see graph of the data? Do you want to see any associated notes (Irena writes a lot of stuff in the wave notes)? Push button "Load data"

Note the "Export multiple data sets?" checkbox. It opens the Multiple Data Export selection panel. This panel enables exporting of many data sets at once. The correct use of this option is to export one data set manually (sets all parameters and export paths), test one data set and then use the Multiple Data set option…. See comments later.



Now we have graph and list of notes. Note, that no attempt is made to create sensible graph. You may have to modify the graph manually if needed.

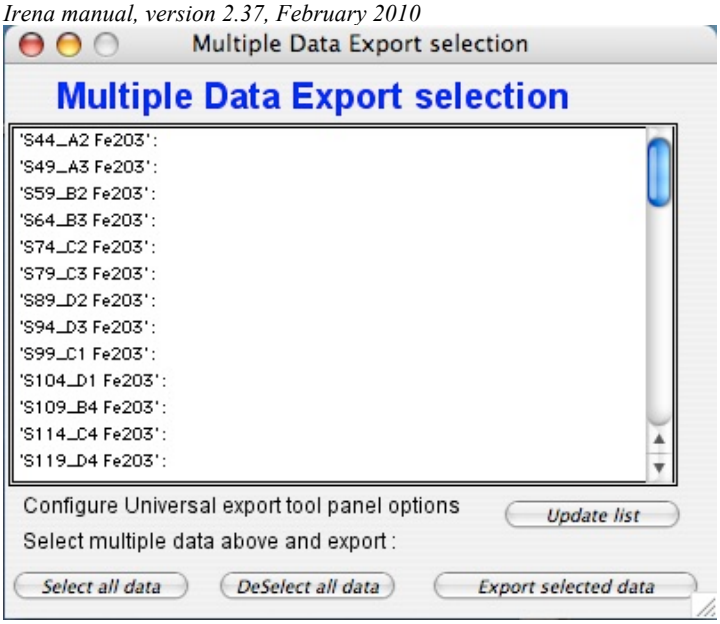
Next select Output options:
Attach notes about data will attach the wave note into the ASCII file. Note, at the bottom of the panel is field where one can insert the separator character (including spaces) if different than default is desired.
Use folder names for output if you are using folder names as anmes for samples, this is sensible…
Use Y wave name for output if your Y wvae name is sample name (e.g., qrs data this type)

"Set export folder" set where to store data. Cannot create folder, create first, then set here. The folder is displayed din red box above the button.

Export file name modify, if default is not good enough
Export file extension set to whatever needed
Header separator change, if different isdesired. Include spaces, if these are desired!!!

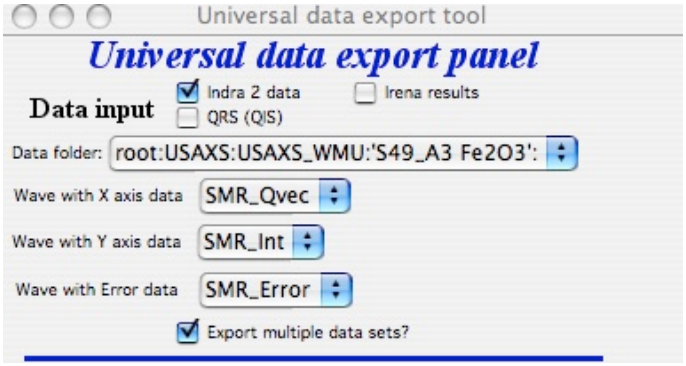
"Export Data & Notes" button does the job. If the data exist, you will be asked if you want to overwrite them.

Multiple data set export option:



There are few items one needs to know about this tool.

1. If you make changes to the main panel, the list of folders in this panel may get stale. Use button "Update list" to update it.
2. There is logic in listing the data which is actually quite complicated… Here are some comments:
    a. The tool started to search for data from parent folder of data selected in the main panel. In the current selection :



The tool start searching from root:USAXS:USAXS_WMU: - if you cannot find your data, select different starting folder in the main panel and update the list. This is to reduce clutter and help users with giant experiments…

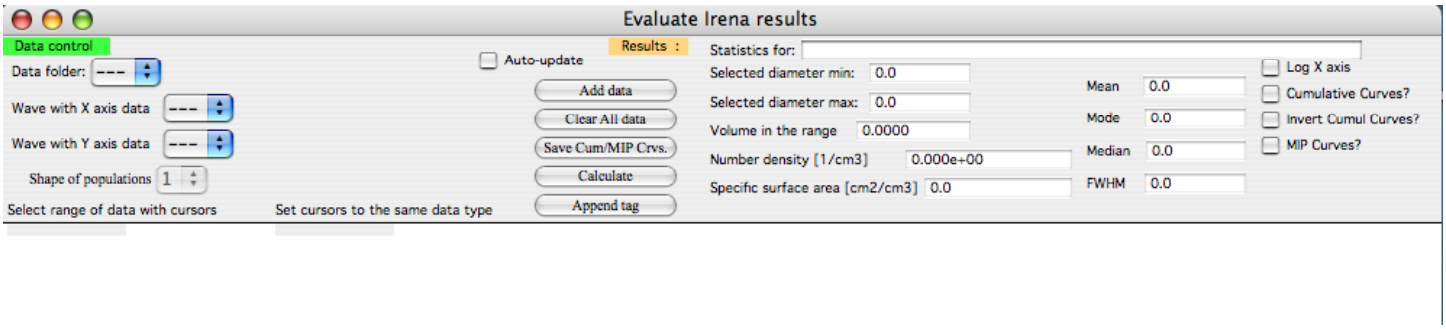
    b. For results… The tool will search for not only the same data tyupe as selected in the main panel, but also same generation! Therefore, if you have in some folders saved multiple results from same tool (you have waves with results like: SizesVolumeDistribution_0, but in some also SizesVolumeDistribution_1, SizesVolumeDistribution_2, etc…) the tool will search only for generation ("_0", "_1",…) selected in the main panel. It just gets really messy to create different logic.

18. Evaluate size distributions tool

This is tool to obtain various details about size distributions – such as results from Size distribution, Modeling I and Modeling II tools. User can obtain mean/mode/median size in range selected by cursors, volume, surface and number of particles (per cm^3) and generate cumulative distributions and even mercury intrusion (MIP) curve (intruded volume with respect to pressure in Psi).
Number of size distributions can be included at one time in the graph, but the graph will become crowded very soon…

18.1 Description

Select "Evaluate size distributions" from the SAS menu.



This tool has all controls in the Control bar at the top of the graph window. For MIP data a new window will be opened. It is also important to follow the history area, as this tool prints important information into there so the user is informed about specific needs or events…

Data selection controls are top left corner. This tool should know all results data from Irena for which it makes sense to be used. If any data type is missing, please let me know and I will add it.
Select data folder and if more than one of useable data types is in that folder, select appropriate data types. You will see only folders, which contain at least one useable data type. You may have to select both X axis data and Y axis data. Very little checking is done on sensibility of the selection here, so be careful.

The usually grayed popup "*Shape of populations*" will be explained below…

Next are buttons with following functions:
"*Auto-update*" checkbox – if selected all data are recalculated when cursors are moved. Note, that cursors have to be on the same data set , or the results all will be NaNs.
"*Add data*" – when data set is selected this button adds the data set into the graph
"*Clear all data*" – will remove all data from graph and clear it.
"*Save Cumul/MIP curves*" – will save new data into the data folder with the original data. Saves the new curves, which exist at that time. Both cumulative size distributions and/or MIP curves. These data can be exported by using the Data export tool.
"*Calculate*" – if the "*Auto-update*" checkbox is not selected, this forces recalculations. Again – if the cursors are not in the graph or not on the same wave, no numbers are calculated!
"*Append tag*" – appends descriptive tag to the graph, so more different data sets can be compared together.

Now the results part:

"*Statistics for*:" … is string with the name of data on which the cursors are and for which the data are calculated.
"*Selected diameter min*" and "*max*" – diameters of current cursor positions so you know where the data are calculated
"*Volume in the range*" – fractional volume of scatterers in the range between cursors. Calculated with correct formula for volume of give form factor used.

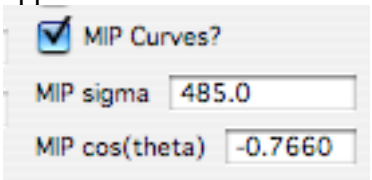
NOTE: for distributions from Modeling I and II it is impossible to decide for the code, where which formula for volume should be used. So if one combines different shapes, there is practically now way one can correctly calculate all of these numbers. Therefore the code will make available the popup "Shape of distributions:" and one can select which shape should be used for the calculations. This is meaningful if the populations are reasonably separate and one knows where which shape dominates. This is problem when one is converting between distributions – so if one is using volume distribution, the volume is correct at all times as there is not conversion needed, but the number of particles may be wrong. If one is using number distribution then number of particles is right but the volume may be wrong. Specific surface area is likely affected all the time, unless one has the right shape. Simply – be careful when using Modeling results with more than one shape of scatterers.

User is informed about need to select right shape by printout in the history area:
"These data may contain mixture of shapes for different populations. Please select the right population number to evaluate"
This is not problem when individual distributions are saved together with the total distribution and evaluated. In such case the code will select correct shape for volume calculations and conversions…

"*Number density*" is number of particles per cm^3 for data between cursors.
"*Specific surfaced area*" is specific surface area between the cursors.

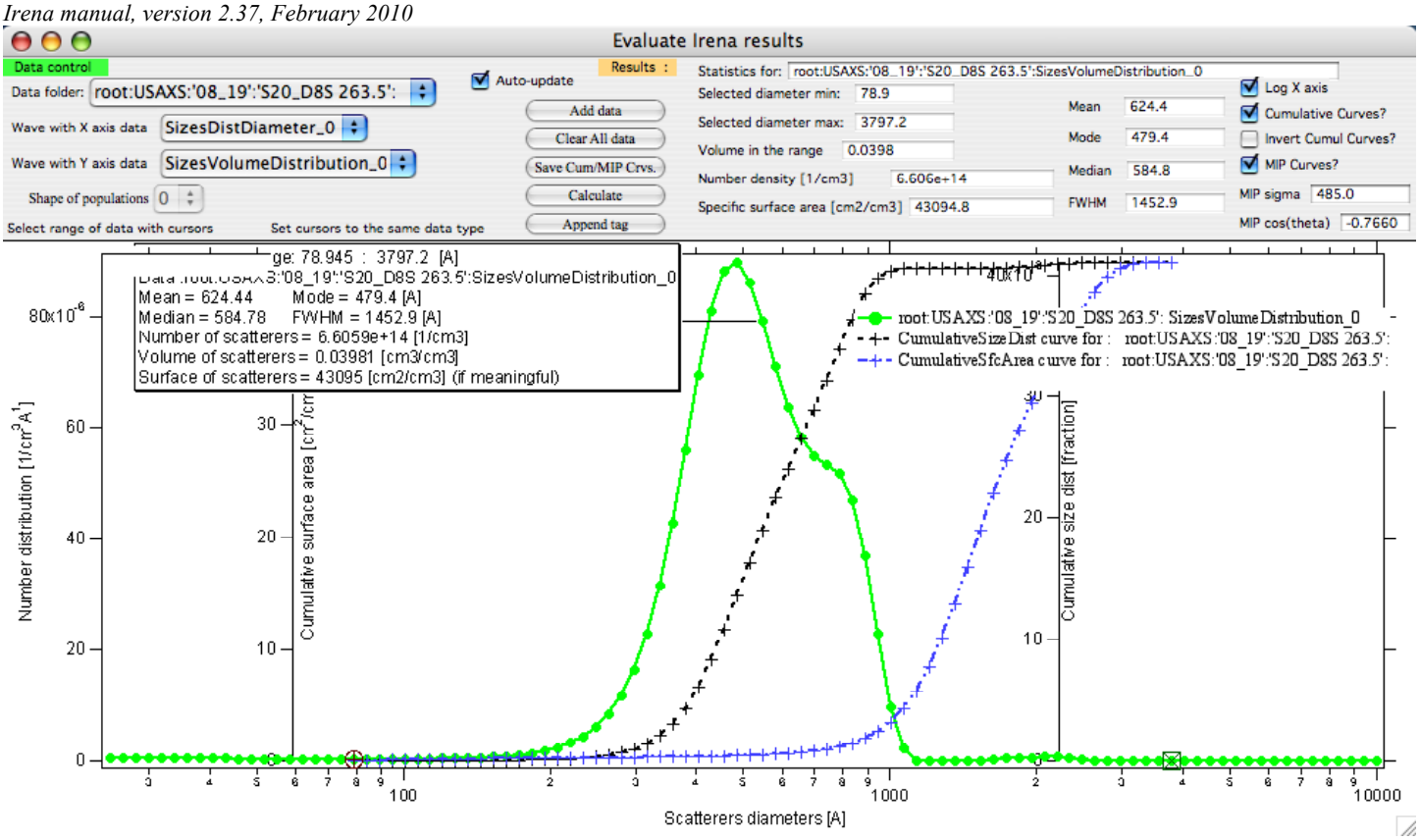
"*Mean*", "*mode*", "*median*" are values evaluated for GIVEN DISTRIBUTION between cursors – evaluated numerically. Note, that of course these will be different for number and volume distributions.
"*FWHM*" is full width at half-maximum value evaluated numerically. This is ONLY meaningful, if the data resemble one single peak. There will always be number there, but it may be not useful if the data are not one single peak.

Last column of checkboxes:
"*Log X*" sets diameter axis (x axis) to log scale.
"*Cumulative curves*" forces calculation of cumulative curves
"*Invert Cumul. Curves*" forces the 0 to be at large sizes. There are some cases when this is useful…
"*MIP curves?"* – if selected MIP curves are calculated and new window with them opens. Few other controls appear also:

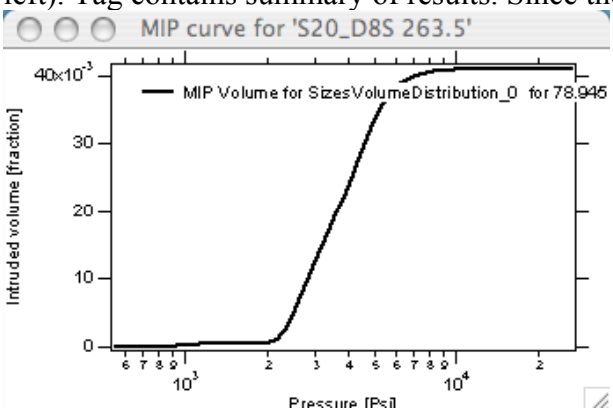


These are two parameters used for MIP calculations. These are generally used values, user can change them if he/she wishes. Sigma is in dynes/cm and cos (theta) is unit less, theta is wetting angle between the material and the mercury…

18.2 Example

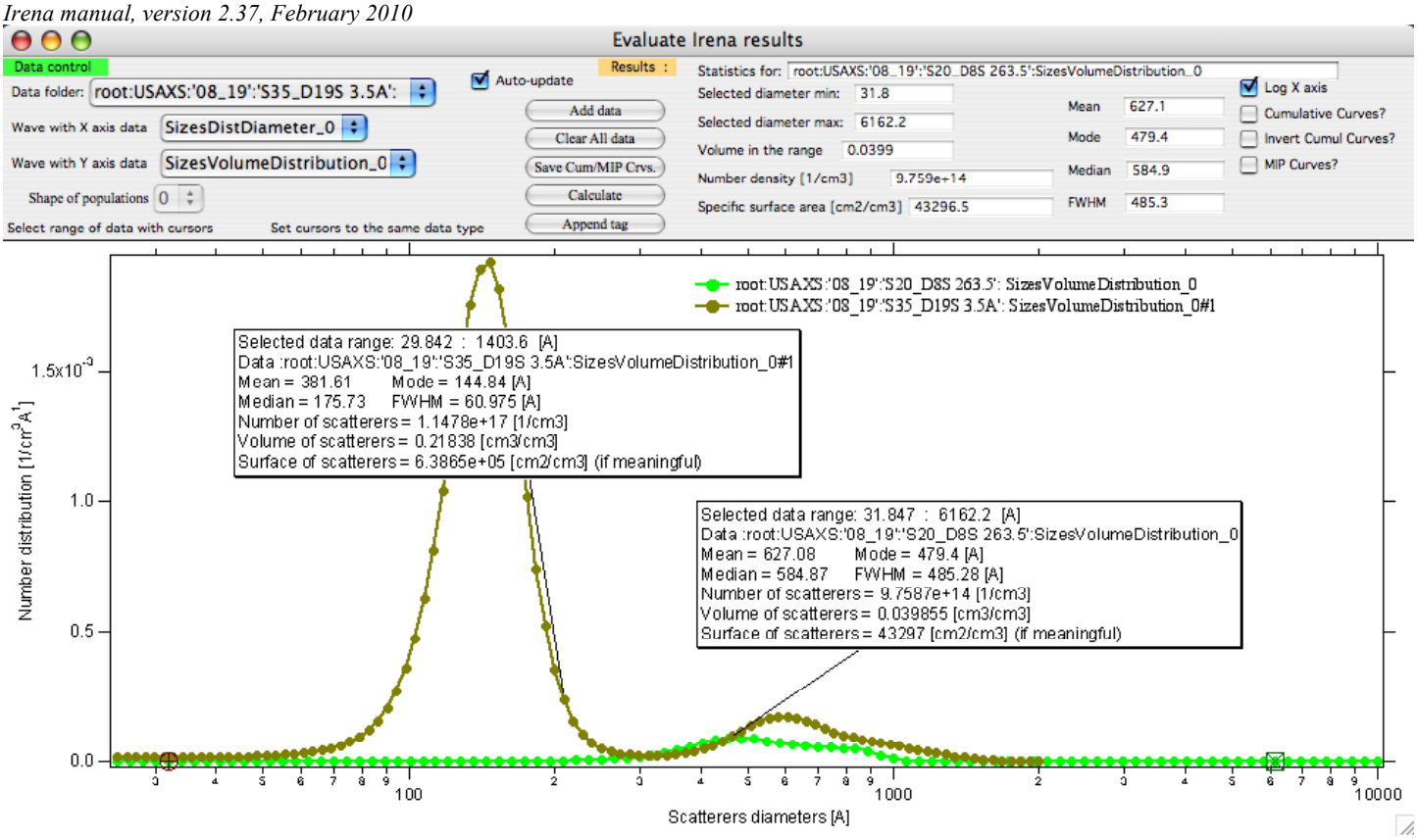


The green data are original data, cursors select the range, which is being evaluated, black curve is Cumulative size distribution volume (it has its own axis on right) and blue is cumulative specific surfaced area (has axis mid left). Tag contains summary of results. Since the MIP curves checkbox was selected, MIP graph was created:



If user chooses to save the new data now, both Cumulative curves as well as MIP curves are going to be saved in the original data folder for export or future use.

Example of comparison of two data sets and use of tags to display results for both:



18.3 New data created

If the user chooses to save the cumulative and/or MIP curves, new data are created in the folder with original size distributions from which these were created. These are named:

MIPVolume_XX
MIPPressure_XX
CumulativeSizeDist_XX
CumulativeSfcArea_Dist_XX
CumulativeDistDiametersDist_XX

With XX being index to guarantee uniqueness.
The index choice is quite complicated and may result in confusion… So here is explanation:

3. First the index of the original data is tested – if Original Data were "SizesVolumeDistribution_2", then the code will test of the index 2 is available. If yes, it will save the data and print result in the history area.
4. If this index is not available, Message is displayed for user and index is increased. User needs to make sure he/she makes note of the right index and keeps notes on this… I have not found more sensible system yet.

When saving data user is informed by printout in the history are what data were created and what generation they were saved in.
Example:
Saved Cumulative data to CumulativeSizeDist_02 / CumulativeSfcArea_Dist_02 / CumulativeDistDiametersDist_02 in folder root:USAXS:'08_19':'S20_D8S 263.5':
Saved MIP data to MIPVolume_01 / MIPPressure_01 in folder root:USAXS:'08_19':'S20_D8S 263.5':

Note, the waves contain descriptive wave notes which can be exported with the ASCII data as header or searched through the Data miner tool.

19 Scripting tool (beta version, December 20, 2007)

This tool enables scripting of (for now) two packages – ***Unified fit and Size Distribution*** – to enable automatic processing of large data sets. During the automatic processing the error messages and most of the user dialogs generated by these two packages are suppressed.

19.1 What this tool does*:*

For each of selected data folders it does following:

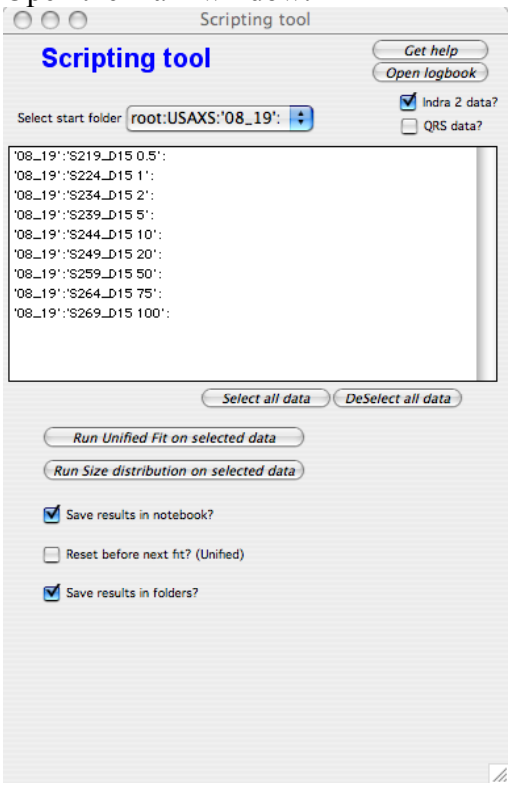
1. Select the data folder and load the data in, create graph (pushes "Graph" button)
2. Sets cursors to the same Q value as previously set
3. Pushes "Fit" button
4. Saves data per selection in the control panel

Limitations: The data selection must be unique for every data set. This requires one data set per folder for qrs data structure. For USAXS data ("Indra2") the data used will be type selected during testing – for example SMR_* or DSM_*, or M_SMR_* etc. These data must exist in all folders. If needed, I can try to update the system and make more options. Send me e-mail if you need more functionality.

19.2 Use of the Scripting tool

Open the tool which you want to script and make a fit on representative data set. Make sure the data range selection (if any) using cursors is correct for all data sets in consideration. Use appropriate data type.

Open the main window:



Select starting folder and from the list of folders here select the ones you want to process. Select type of data to be processed… Important: The data type should really be same as the test case used to setup the parameters of the fits (see below)….

Note, that the order here will be from top to bottom. There is at this time no option to change the order. Let me know if ordering should be done. Seems complicated, as there are not simple Igor tools for this.
The buttons at the top:
"Get help" creates panel with short help for this tool.
"Open logbook" opens one of two logbooks… If this tool was used to create its own notebook for results, this one will open. This notebook contains selected results and also graphs. If this notebook does not exist, standard internal logbook will be opened. This one contains similar data, but no graphs.

Select options:
1. *Save results in notebook* – will create notebook in which graph and summary of results for each data sets will be printed.
2. *Reset before next run* – valid for Unified. Will reset to original values from test case after each fit. This is useful when the changes are not progressive but more or less random and previous result could be worse starting point that the test case. If not selected, the last result is used as starting model for next sample.
3. *Save results in folders* – will copy usual results waves in the original data folders so they can be plotted or looked at again.

Note:
Without "Save results in notebook" and/or "Save results in folders" there will be no output from the procedure and you will not learn anything…

To run:
Select representative case example (and test extreme cases) and run the tool on this example data. Make sure all parameters are set correctly. This includes:
Proper selection of fit checkboxes and limits on Unified…
Proper error weighing and background value in Size distribution.
Proper range of data selection with cursors (if applicable) for both tools.
When ready push button "Run …. on selected data"

Final comments

This manual is apparently never ever finished… Real apologies to everyone for this, but it seems to be impossible to keep up with changes and modifications.